(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 September 2001 (13.09.2001)

PCT

(10) International Publication Number
WO 01/66739 A1

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, C12N 5/10, 15/62, C07K 16/18

(21) International Application Number: PCT/US01/04475

(22) International Filing Date: 12 February 2001 (12.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/187,572	7 March 2000 (07.03.2000)	US
60/187,848	8 March 2000 (08.03.2000)	US

(71) Applicant *(for all designated States except US)*: **THE GOVERNMENT OF THE UNITED STATES OF AMERICA as represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES** [US/US]; 6011 Executive Boulevard, Suite 325, Rockville, MD 20852 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **THORGEIRSSON, Snorri, S.** [US/US]; 9305 Kingsley Avenue, Bethesda, MD 20814 (US). **SANDERS, Sean** [ZA/US]; Apartment 202, 10308 Rockville Pike, Rockville, MD 20852-3314 (US).

(74) Agent: **DAIGNAULT, Ronald, A.**; Merchant & Gould P.C., P.O. Box 2903, Minneapolis, MN 55402-0903 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/66739 A1

(54) Title: çDNA ENCODING A TRAG GENE (TGF- β RESISTANCE ASSOCIATED GENE) AND ITS PROTEIN PRODUCT

(57) Abstract: Nucleic acids that encode novel polypeptides, designated in the present application as "TRAG" (TGF-β Resistance Associated Gene) are provided. TRAG is overexpressed in a number of transformed rat liver epithelial (RLE) cell lines resistant to the growth inhibitory effect of TGF-β1 as well as in primary liver tumors. Compositions including TRAG chimeras, nucleic acids encoding TRAG, and antibodies to TRAG are also provided. Methods of using TRAG to screen for cancer cells, to screen for aggressive metastasis in cells, to screen for targets for cancer therapy, and to study cell proliferation are further provided.

cDNA ENCODING A TRAG GENE (TGF- β RESISTANCE ASSOCIATED GENE) AND ITS PROTEIN PRODUCT

This application is being filed as a PCT International Patent application on February 12, 2001, designating all countries, in the name of The Government of the United States, represented by the Secretary, Department of Health and Human Services (applicant for all countries except the U.S.), and in the names of Snorri S. Thorgeirsson, a U.S. citizen, and Sean Sanders, a South African citizen (applicants for U.S. only).

FIELD OF THE INVENTION

The present invention relates generally to the isolation and characterization of novel DNA and polypeptides, designated herein as "TRAG."

BACKGROUND OF THE INVENTION

The ligand transforming growth factor–beta (TGF–β) has a wide range of physiological and pathological effects in epithelial cells (which includes liver cells), including inhibition of cell proliferation, stimulation of cell differentiation, matrix production, and apoptosis. (J. Massagué, Annual Rev. Cell Biol., 6:597 (1990)). TGF–β can suppress uncontrolled cell proliferation and induce cell death (apoptosis) in damaged/redundant cells by signaling through membrane receptors (TGF–β receptor type I, TβRI; TGF–β receptor type II, TβRII) and intracellular signal transduction proteins (for example, SMAD2, SMAD3, and SMAD4). (M. Kawabata et al., J. Biochem., 125:9 (1999)). More importantly, many tumor types, and particularly those in the liver, have been shown to lose sensitivity to TGF–β. That is, these cells are not susceptible to abrogation of cell division or death induced by TGF–β. This can occur in a number of different ways, including through loss of one of the two TGF–β receptors or through disruption of the intracellular signaling pathway (such as loss of mutation of one of the SMAD proteins). (See, for example, K.R. Cho et al., J. Biol. Chem. Suppl., 137 (1992) and A. Kiss et al., Clin. Cancer Res., 3:1059 (1997)).

Recently, a number of new proteins that interact with one or more components of the TGF–β signaling pathway have been discovered. These include TRIP–1, which has been shown to associate with, and be biochemically modified by, TβRII, and STRAP, which can interact with both TβRI and TβRII. (R.H. Chen et al., Nature, 377:548 (1995)

and P.K. Datta et al., J. Biol. Chem., 273(52):34671 (1998)). Both of these proteins contain conserved domains known as WD repeats. These repeats are ~40 amino acids in length and follow a loosely conserved consensus sequence. (E.J. Neer et al., Nature, 371:297 (1994)). WD repeat–containing proteins are believed to interact through these conserved domains and are involved in a plethora of cellular and molecular pathways including signal transduction, gene regulation, protein trafficking, and RNA processing. (E.J. Neer et al., Nature, 371:297 (1994)).

It is desirable to identify and characterize new cellular genes encoding proteins involved in TGF–β signaling. The identification and characterization of new genes encoding proteins involved in TGF–β signaling are particularly useful for better understanding TGF–β signaling pathways such as, for example, useful for studying mechanisms of cell proliferation and the means to modulate such activity, and particularly useful for screening targets for cancer therapy.

SUMMARY OF THE INVENTION

Applicants have identified nucleotide sequences that encode a novel polypeptide, designated in the present application as "TRAG" (TGF–β Resistance Associated Gene), which exhibits a number of characteristics that make it a useful tool for studying cell–cycle control and oncogenesis. TRAG is a previously undescribed gene. TRAG has multiple WD repeat elements and two tyrosine phosphorylation motifs. WD repeat elements are known to be involved in protein–protein interactions, which control many aspects of cell growth, differentiation, and survival. Tyrosine phosphorylation motifs can play a role in the modulation of protein function. As such, this novel protein has a variety of applications in the identification, characterization, and regulation of activities associated with cellular regulation as well as processes associated with oncogenesis. It is believed that TRAG plays an important role in TGF–β signal transduction, normal cell function, and organogenesis.

In one embodiment, the invention provides an isolated nucleic acid molecule that includes nucleotides that encode a TRAG polypeptide. For example, the isolated nucleic acid can include DNA encoding a TRAG polypeptide having amino acid residues 1 to 1488 of Table 1, amino acid residues 1 to 1490 of Table 2, or amino acid residues 1 to 1489 of Table 3 or can include DNA complementary to such an encoding nucleic acid sequence of Table 1, 2, or 3, which remains stably bound to it under at least moderate,

and optionally, under high stringency conditions. In another embodiment, the invention provides a vector comprising a gene encoding a TRAG polypeptide. A host cell comprising such a vector is also provided. By way of example, the host cells may be *E. coli*, yeast, insect, fungal, or mammalian cells. A process for producing TRAG polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of TRAG. If desired, the TRAG polypeptide can be recovered.

In one embodiment, the invention provides an isolated polypeptide comprising a TRAG fragment, wherein the TRAG fragment comprises multiple WD repeat elements and two tyrosine phosphorylation motifs. In a favored embodiment, the isolated polypeptide exhibits TRAG–like activity and, typically, is capable of interacting with other WD repeat element–containing proteins. In another embodiment, the invention provides isolated TRAG polypeptide. In particular, the invention provides isolated native sequence TRAG polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 1488 of Table 1, an amino acid sequence comprising residues 1 to 1490 of Table 2, or an amino acid sequence comprising residues 1 to 1489 of Table 3. In a related embodiment, the invention provides chimeric molecules comprising TRAG polypeptide fused to a heterologous polypeptide or amino acid sequence. An example of such a chimeric molecule is a factor that includes a TRAG fused to a protein such as the maltose–binding protein. In yet another embodiment, the invention provides a polypeptide capable of specifically binding a TRAG polypeptide such as an antibody specific for a TRAG polypeptide.

In other embodiments, the invention provides methods for using TRAG–RELATED polypeptides and nucleic acids for studying and modulating mechanisms involved in cellular proliferation. In one embodiment, the invention provides a method of modulating cellular phenotype by controlling the level of TRAG expression within the cell. For example, mammalian cells can be transfected with a DNA vector encoding a TRAG polypeptide having the amino acid sequence of Table 1, 2, or 3. The TRAG polypeptide can be expressed in the cells, and cells having an altered phenotype such as, for example, cancer cells or cells having aggressive metastasis, can then be selected.

Alternatively, the invention provides a method of reducing TRAG expression via antisense oligonucleotides to effect a cellular phenotype such as TGF–β sensitivity. In a related embodiment, the invention provides methods for effecting the interaction between TRAG and other WD repeat element–containing proteins.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A shows the normal expression pattern of TRAG in various rat tissues according to a Northern blot analysis. This Northern blot shows mRNA (total and poly-A^+) extracted from 3 rat cell lines (RLE, B5T, and C4T).

Figure 1B shows a graphical representation of the data from Figure 1A after the levels of TRAG mRNA were quantified. Data was quantified using a densimetric scan of the phosphoimager plate (to which the radioactive membrane was exposed) by making use of ImageQuant™ Version 3.3 software (Molecular Dynamics).

Figure 2A shows the genetic localization of the TRAG gene to chromosome 18 in mouse (region is 18D.1-E.3). This localization was determined by using a DNA probe complementary to a portion of the TRAG gene, covering exons 14 to 21 (left panel). The right panel shows a similar experiment, but performed using a DNA probe that reacts exclusively with the entire mouse chromosome 18.

Figure 2B shows the genetic localization of the TRAG gene to chromosome 18 in human (region is 18q 21.1-22). This localization was determined by using a ~5kb DNA probe against a centrally located region of the TRAG gene spanning exons 12 to 14 (left panel). The right panel shows a similar experiment, but performed using a DNA probe that reacts exclusively with the entire human chromosome 18.

Figure 3A shows the size and qualitative levels of TRAG mRNA expressed in normal rat tissue from heart, brain, spleen, lung, liver, muscle, kidney, and testis. mRNA transcripts having sizes of approximately 7.2 kb and 3.5 kb were identified. The brain tissue shows a smear of multiple mRNAs, which can be the result of alternative splicing of a single mRNA transcript.

Figure 3B shows a dot blot hybridization using a commercial membrane containing rat mRNA from brain, heart, gastrointestinal tract (GIT), numerous internal organs including liver and testis (indicated), tumor cell lines (labeled "lines"), and fetal tissue (from top: brain, heart, kidney, liver, spleen, thymus, and lung). TRAG mRNA was expressed in all tissues but appeared to be most abundant in brain.

Figure 3C shows a dot blot hybridization using a commercial membrane containing mouse mRNA from (top, L to R) brain, eye, liver, lung, kidney, heart, skeletal muscle, smooth muscle, pancreas, thyroid, thymus, submaxillary gland, spleen, testis, ovary, prostate, epididymus, uterus, and embryo (7 days, 11 days, 15 days, and 17 days). TRAG expression was identified in all tissues.

Figure 4 shows a Western blot of TRAG protein extracted from three rat cell lines (B5T, C4T, and RLE phi 13). This blot shows elevated levels of TRAG protein in transformed B5T and C4T cells relative to untransformed RLE cells.

Figure 5A shows a Western blot of TRAG protein extracted from chemically transformed rat cell lines (AFL–B8 and AFL–D8) and virally transformed rat cell lines (3611T2, 3611T5, and J2–14). This blot also includes samples from RLE and B5T cells as negative and positive controls, respectively. A Northern blot, which shows the expression of TRAG mRNA in these cell lines, is also shown.

Figure 5B shows a Western blot of TRAG protein extracted from tumor cell lines derived from double transgenic c–*myc*/TGF–α mice. A Northern blot, which shows the TRAG mRNA levels in these cell lines, is also shown.

Figure 5C shows a Western blot of TRAG protein extracted from Alex, Chang, FOCUS, HepG2, Huh–7, Sk–Hep–1, WRL–68, and HeLa human tumor cell lines. A Northern blot, which shows the TRAG mRNA levels in these cell lines, is also shown.

Figure 5D shows a Western blot of TRAG protein extracted from primary tumors taken from double transgenic c–*myc*/TGF–α mice. This blot also includes samples from RLE and B5T cells as negative and positive controls, respectively.

Figure 6A shows confocal microscope images of RLE and B5T cells transfected with either a control green fluorescent protein (GFP only) plasmid or with a vector containing the TRAG gene fused to the green fluorescent protein (TRAG–GFP). The images show that the TRAG protein is localized to the cytoplasm of both RLE and B5T cells.

Figure 6B shows an immunohistochemical stain of paraformaldehyde–fixed RLE and B5T cells. The TRAG protein was stained by using an anti–TRAG polyclonal antibody. The stain shows that the TRAG protein is localized to the cytoplasm of both RLE and B5T cells and appears to be perinuclear.

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

The terms "TRAG polypeptide" and "TRAG" when used herein encompass native sequence TRAG and TRAG variants (which are further defined herein). TRAG may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

The TRAG polypeptide, which may be a fragment of a native sequence, contains multiple WD repeat elements and 2 tyrosine phosphorylation motifs.

A "native sequence TRAG" is a polypeptide having the same amino acid sequence as TRAG derived from nature. Such native sequence TRAG can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence TRAG" specifically encompasses naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of the TRAG. In one embodiment of the invention, the native sequence TRAG is a mature or full-length native sequence TRAG polypeptide comprising amino acids 1 to 1488 of Table 1. Alternately, the TRAG polypeptide comprises amino acid residues 1 to 1490 of Table 2 or amino acid residues 1 to 1489 of Table 3.

"TRAG variant" means a functionally active TRAG as defined below having at least about 80% amino acid sequence identity with TRAG, such as the TRAG polypeptide having the deduced amino acid sequence shown in Tables 1, 2, or 3 for a full-length native sequence TRAG. Such TRAG variants include, for instance, TRAG polypeptides wherein one or more amino acid residues are added to or deleted from the N-terminus or C-terminus of the sequence of Tables 1, 2, or 3. Ordinarily, a TRAG variant will have at least about 80% or 85% amino acid sequence identity with native TRAG sequences, more preferably at least about 90% amino acid sequence identity. Even more preferably a TRAG variant will have at least about 95% amino acid sequence identity with the native TRAG sequence of Tables 1, 2, or 3. As noted above, TRAG variants include multiple WD repeat elements and 2 tyrosine phosphorylation motifs. Functionally active TRAG variants typically have at least about 50 amino acid residues and preferably at least about 100 amino acid residues.

"Percent (%) amino acid sequence identity" with respect to the TRAG sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the TRAG sequence after aligning the sequences in the same reading frame and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full

length of the sequence being compared.

"Percent (%) nucleic acid sequence identity" with respect to the TRAG sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the TRAG sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising TRAG, or a functional fragment thereof, fused to a "tag polypeptide." The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, or which can be identified by some other agent, yet is short enough that it does not interfere with the activity of TRAG. The tag polypeptide preferably also is sufficiently unique so that the antibody does not substantially cross–react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and about 50 amino acid residues (preferably, between about 10 and about 20 residues).

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the polypeptide will be purified to a degree sufficient to obtain N–terminal or internal amino acid sequence by use of a spinning cup sequenator or to homogeneity by SDS–PAGE under nonreducing or reducing conditions using Coomassie blue or silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells because at least one component of the TRAG natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step (referred to herein as an "isolated and purified polypeptide").

An "isolated" TRAG nucleic acid molecule is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which

is ordinarily associated in the natural source of the TRAG nucleic acid. An isolated TRAG nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated TRAG nucleic acid molecules therefore are distinguished from the TRAG nucleic acid molecule as it exists in natural cells. However, an isolated TRAG nucleic acid molecule includes TRAG nucleic acid molecules contained in cells that ordinarily express TRAG where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking may be accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers may be used in accordance with conventional practice.

"Polynucleotide" and "nucleic acid" refer to single- or double-stranded molecules, which may be DNA, comprised of the nucleotide bases A, T, C and G, or RNA, comprised of the bases A, U (substitutes for T) , C, and G. The polynucleotide may represent a coding strand or its complement. Polynucleotide molecules may be identical in sequence to the sequence that is naturally occurring or may include alternative codons that encode the same amino acid as that which is found in the naturally occurring sequence (*See* Lewin, Genes V, Oxford University Press, Chapter 7, pp. 171-174 (1994)). Furthermore, polynucleotide molecules may include codons that represent

conservative substitutions of amino acids as described. The polynucleotide may represent genomic DNA or cDNA.

"Polypeptide" refers to a molecule comprised of amino acids that correspond to those encoded by a polynucleotide sequence that is naturally occurring. The polypeptide may include conservative substitutions where the naturally occurring amino acid is replaced by one having similar properties, where such conservative substitutions do not alter the function of the polypeptide (*See* Lewin, Genes V, Oxford University Press, Chapter 1, pp. 9-13 (1994)).

The term "antibody" is used in the broadest sense and specifically covers single anti-TRAG monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies) and anti-TRAG antibody compositions with polyepitopic specificity. The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, *i.e.*, the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

The term "mammal" as used herein refers to any mammal classified as a mammal, including humans, cows, horses, dogs, and cats. In a preferred embodiment of the invention, the mammal is a human.

II. Compositions and Methods of the Invention

A. TRAG Nucleic Acids and Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as TRAG. In particular, Applicants have identified and isolated genes and cDNA encoding TRAG polypeptides, as disclosed in further detail in the Examples below. Using sequence homology searches, Applicants found that TRAG (as shown in Tables 1, 2, and 3) contains multiple WD repeat elements, which are found in other proteins such as TRIP-1 and STRAP, known to interact with proteins in the TGF-β signaling pathway. Two tyrosine phosphorylation motifs were also identified. These motifs can play a role in the modulation of protein function through addition or removal of phosphate molecules at the tyrosine residue.

In addition to the full-length native sequence TRAG and soluble forms of TRAG

described herein, it is contemplated that TRAG variants can be prepared. TRAG variants can be prepared by introducing appropriate nucleotide changes into the TRAG nucleotide sequence, or by synthesis of the desired TRAG polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the TRAG, such as changing the number or position of glycosylation sites or altering the protein binding characteristics. Variations in the native full-length sequence TRAG or in various domains of the TRAG described herein, can be made, for example, using any of the techniques and guidelines for conservative and nonconservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. For example, amino acid substitutions within the tyrosine phosphorylation motifs are contemplated, such as conservative substitutions at one or both of these residues.

Variations may be a substitution, deletion, or insertion of one or more codons encoding the TRAG that results in a change in the amino acid sequence of the TRAG as compared with the native sequence TRAG. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the TRAG. Guidance in determining which amino acid residue may be inserted, substituted, or deleted without adversely affecting the desired activity may be found by comparing the sequence of the TRAG with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity in any of the *in vitro* assays described in the Examples below.

It is a well-established principle of protein chemistry that certain amino acid substitutions, entitled "conservative amino acid substitutions," can frequently be made in a protein without altering either the conformation or the function of the protein. Such changes include substituting any of isoleucine (I), valine (V), and leucine (L) for any other of these hydrophobic amino acids; aspartic acid (D) for glutamic acid (E) and vice versa; glutamine (Q) for asparagine (N) and vice versa; and serine (S) for threonine (T) and vice versa. Other substitutions can also be considered conservative, depending on the environment of the particular amino acid and its role in the three-dimensional structure of

the protein. For example, glycine (G) and alanine (A) can frequently be interchangeable, as can alanine and valine (V).

Methionine (M), which is relatively hydrophobic, can frequently be interchanged with leucine and isoleucine, and sometimes with valine. Lysine (K) and arginine (R) are frequently interchangeable in locations in which the significant feature of the amino acid residue is its charge and the differing pK's of these two amino acid residues are not significant. Still other changes can be considered "conservative" in particular environments.

Variations can be made using methods known in the art such as site–directed mutagenesis, alanine scanning, and PCR mutagenesis. Site–directed mutagenesis (Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)), cassette (Wells et al., Gene, 34:315 (1985)), restriction selection mutagenesis (Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)) or other known techniques can be performed on the cloned DNA to produce the TRAG variant DNA. Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side–chain beyond the beta–carbon and is less likely to alter the main–chain conformation of the variant. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions (Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)). If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

As discussed above, redundancy in the genetic code permits variation in TRAG gene sequences. In particular, one skilled in the art will recognize specific codon preferences by a specific host species and can adapt the disclosed sequence as preferred for a desired host. For example, preferred codon sequences typically have rare codons (i.e., codons having a useage frequency of less than about 20% in known sequences of the desired host) replaced with higher frequency codons. Codon preferences for a specific organism may be calculated, for example, by utilizing codon usage tables available on the INTERNET at the following address: http://www.dna.affrc.go.jp/~nakamura/codon.html. Nucleotide sequences that have been optimized for a particular host species by replacing any codons having a useage

frequency of less than about 20% are referred to herein as "codon optimized sequences."

Additional sequence modifications are known to enhance protein expression in a cellular host. These include elimination of sequences encoding spurious polyadenylation signals, exon/intron splice site signals, transposon–like repeats, and/or other such well–characterized sequences that may be deleterious to gene expression. The GC content of the sequence may be adjusted to levels average for a given cellular host, as calculated by reference to known genes expressed in the host cell. Where possible, the sequence may also be modified to avoid predicted hairpin secondary mRNA structures. Other useful modifications include the addition of a translational initiation consensus sequence at the start of the open reading frame, as described in Kozak, Mol. Cell Biol., 9:5073–5080 (1989). Nucleotide sequences that have been optimized for expression in a given host species by elimination of spurious polyadenylation sequences, elimination of exon/intron splicing signals, elimination of transposon–like repeats, and/or optimization of GC content in addition to codon optimization are referred to herein as an "expression enhanced sequence."

B. Modifications of TRAG

Covalent modifications of TRAG are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of the TRAG with an organic derivatizing agent that is capable of reacting with selected side chains or the N– or C– terminal residues of the TRAG. Derivatization with bifunctional agents is useful, for instance, for crosslinking TRAG to a water–insoluble support matrix or surface for use in the method for purifying anti–TRAG antibodies, and vice–versa. Commonly used crosslinking agents include, *e.g.*, 1,1–bis(diazoacetyl)–2–phenylethane, glutaraldehyde, N–hydroxysuccinimide esters, for example, esters with 4–azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'–dithiobis(succinimidylpropionate), bifunctional maleimides such as bis–N–maleimido–1,8–octane and agents such as methyl–3–[(p–azidophenyl)dithio]propioimidate. In the alternative, TRAG can be joined to a detectable label such as a radioactive isotope such as I^{125} or P^{32}, an enzyme such as horseradish peroxidase or alkaline phosphatase, a fluorophore such as fluorescein isothiocyanate or a chromophore (Current Protocols In Molecular Biology, Volume 2, Units 10, 11 and 14, Frederick M. Ausubul et al. eds., 1995; Molecular Cloning, A Laboratory Manual, § 12, Tom Maniatis et al. eds., 2d ed.

1989).

Another type of covalent modification of the TRAG polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence TRAG, and/or adding one or more glycosylation sites that are not present in the native sequence TRAG. Addition of glycosylation sites to the TRAG polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence TRAG (for O–linked glycosylation sites). The TRAG amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the TRAG polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids. Another means of increasing the number of carbohydrate moieties on the TRAG polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259–306 (1981).

The TRAG of the present invention may also be modified in a way to form a chimeric molecule comprising TRAG fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of the TRAG with a tag polypeptide which provides an epitope to which an anti–tag antibody can selectively bind. The epitope tag is generally placed at the amino– or carboxyl– terminus of the TRAG. The presence of such epitope–tagged forms of the TRAG can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the TRAG to be readily purified by affinity purification using an anti–tag antibody or another type of affinity matrix that binds to the epitope tag.

In an alternative embodiment, the chimeric molecule may comprise a fusion of the TRAG with an immunoglobulin or a particular region of an immunoglobulin. The TRAG may be fused to any one of a variety of known fusion protein partners that are well known in the art such as maltose binding protein, *LacZ*, thioredoxin, or an immunoglobulin constant region (Current Protocols In Molecular Biology, Volume 2, Unit 16, Frederick M. Ausubul et al. eds., 1995; Linsley et al., J.Exp. Med., 174:561-566 (1991)). In a preferred embodiment, this fusion partner is a non-TRAG binding molecule

so as to prevent difficulties associated with intramolecular interactions. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule. Other fusion proteins and tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 (Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7, and 9E10 antibodies thereto (Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky et al., Protein Engineering, 3(6):547-553 (1990)). Other tag polypeptides include the Flag-peptide (Hopp et al., BioTechnology, 6:1204-1210 (1988)); the KT3 epitope peptide (Martin et al., Science, 255:192-194 (1992)); an I-tubulin epitope peptide (Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)).

C. Preparation of TRAG

The description below relates primarily to production of TRAG by culturing cells transformed or transfected with a vector containing TRAG nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare TRAG. For instance, the TRAG sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques (*see, e.g.*, Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)). *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the TRAG may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length TRAG.

1. Isolation of DNA Encoding TRAG

Due to the large size of TRAG mRNA, and therefore cDNA, both the rat and mouse TRAG genes (coding regions only) were isolated using the RT-PCR and RACE-PCR techniques, as is well-known in the art. Primer oligonucleotides were synthesized (Gibco, BRL Life Technologies) and used to amplify cDNA (produced directly from rat

or mouse mRNA), as described in Example 1. Three portions of the mouse TRAG gene so amplified were joined through the use of restriction digestion and subcloning, as described in Example 8, to yield a full length gene of 4,467 nucleotides of coding sequence.

Alternately, DNA encoding TRAG may also be obtained from a cDNA library prepared from tissue expressing a TRAG mRNA. Accordingly, human TRAG DNA can be conveniently obtained from a cDNA library prepared from human tissue. The TRAG–encoding gene may also be obtained from a genomic library or by oligonucleotide synthesis. Libraries can be screened with probes (such as antibodies to the TRAG or oligonucleotides of at least about 20–80 bases) designed to identify the gene of interest or the protein encoded by it. Illustrative libraries include mouse kidney cDNA library (mouse kidney 5'-stretch cDNA, Clonetech laboratories, Inc.) and human liver cDNA library (human liver 5' stretch plus cDNA, Clonetech Laboratories, Inc.). Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding TRAG is to use PCR methodology (Sambrook et al., supra; Dieffenbach et al., PCR Primer:A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)).

For cDNA library screening, oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art and include the use of radiolabels like ^{32}P–labeled ATP, biotinylation, or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full–length sequence can be determined through sequence alignment using computer software programs which employ various algorithms to measure homology.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed

herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for TRAG production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH, and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of transfection are known to the ordinarily skilled artisan, for example by using lipofectin, $CaPO_4$, or electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes or other cells that contain substantial cell-wall barriers. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transformations have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, *e.g.*, polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are

not limited to eubacteria, such as Gram–negative or Gram–positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635)

Suitable host cells for the expression of glycosylated TRAG are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as Drosophila S2 and Spodoptera Sf9 cells. See e.g. Current Protocols In Molecular Biology, Volume I, Unit 16, Frederick M. Ausubul et al. eds., 1995. Examples of useful mammalian host cell lines include rat liver epithelial cells, Hugget, A. C. et. al., supra, Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS–7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen Virol., 36:59 (1977)); Chinese hamster ovary cells/–DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243–251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (*e.g.*, cDNA or genomic DNA) encoding TRAG may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector

by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques that are known to the skilled artisan.

TRAG may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N–terminus of the mature protein or

polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the TRAG DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat–stable enterotoxin II leaders. For yeast secretion the signal sequence may be, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α–factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram–negative bacteria, the P2 plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, *e.g.*, the gene encoding D–alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the TRAG nucleic acid, such as Neomycin, DHFR , or thymidine kinase. An appropriate host cell when wild–type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp*1 gene present in the yeast plasmid YRp7 (Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)). The *trp*1 gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4–1 (Jones, Genetics, 85:12 (1977)).

Expression and cloning vectors usually contain a promoter operably linked to the TRAG nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a

variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β–lactamase and lactose promoter systems (Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)), alkaline phosphatase, a tryptophan (trp) promoter system (Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776), and hybrid promoters such as the tac promoter (deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21–25 (1983)). Promoters for use in bacterial systems also will contain a Shine–Dalgarno (S.D.) sequence operably linked to the DNA encoding TRAG.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3–phosphoglycerate kinase (Hitzeman et al., J. Biol. Chem., 255:2073 (1980)) or other glycolytic enzymes (Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)), such as enolase, glyceraldehyde–3–phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose–6–phosphate isomerase, 3–phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde–3–phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

TRAG transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis–B virus, and Simian Virus 40 (SV40), from heterologous mammalian promoters, *e.g.*, the actin promoter or an immunoglobulin promoter, and from heat–shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding TRAG by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis–acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, I –fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100–270), the cytomegalovirus early promoter

enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the TRAG coding sequence but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding TRAG. Still other methods, vectors, and host cells suitable for adaptation to the synthesis of TRAG in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620–625 (1981); Mantei et al., Nature, 281:40–46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to determine the transcription of mRNA (Thomas, Proc. Natl. Acad. Sci. USA, 77:5201–5205 (1980)), dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA–RNA hybrid duplexes or DNA–protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence TRAG polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to TRAG DNA and encoding a specific antibody epitope.

5. Purification of Polypeptide

Forms of TRAG may be recovered from culture medium or from host cell lysates. Cells employed in expression of TRAG can be disrupted by various physical or chemical means, such as freeze–thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify TRAG from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion–exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation–exchange resin such as DEAE; chromatofocusing; SDS–PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G–75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope–tagged forms of the TRAG. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer–Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular TRAG produced.

D. Uses for TRAG

Nucleotide sequences (or their complement) encoding TRAG have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping, and in the generation of antisense RNA and DNA. TRAG nucleic acid will also be useful for the preparation of TRAG polypeptides by the recombinant techniques described herein. TRAG polypeptides have various applications in the art, including uses for evaluating factors that interact with and/or control TGF-β signaling as means for understanding both cell proliferation control and oncogenesis. Moreover, TRAG genes may be introduced into cells to effect mechanisms mediated by TGF-β as well as processes involved in oncogenesis.

1. Screening Methods Utilizing TRAG Nucleic Acids.

The full–length native sequence TRAG (Tables 1, 2, and 3) gene, or portions thereof, may be used as hybridization probes for a cDNA library to isolate, for instance, still other genes (like those encoding naturally–occurring variants of TRAG or TRAG

from other species) that have a desired sequence identity to the TRAG sequences disclosed in Tables 1, 2, or 3. Optionally, the length of the probes will be about 20 to about 500 bases. The hybridization probes may be derived from the nucleotide sequence or from genomic sequences including promoters, enhancer elements, and introns of native sequence TRAG. By way of example, a screening method will include isolating the coding region of the TRAG gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the TRAG gene of the present invention can be used to screen libraries of human cDNA, genomic DNA, or mRNA to determine to which members of such libraries the probe hybridizes. Hybridization techniques are described in further detail in the Examples below.

Nucleotide sequences encoding a TRAG can also be used to construct hybridization probes for mapping the gene that encodes that TRAG and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

Screening assays can be designed to find lead compounds that mimic the biological activity of a native TRAG or a ligand or receptor for TRAG. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein–protein binding assays, biochemical screening assays, immunoassays, and cell–based assays, which are well characterized in the art.

2. Modulation of TRAG protein Expression via TRAG Antisense Oligonucleotides.

Antisense technology entails the administration of exogenous oligonucleotides that bind to a target polynucleotide located within the cells. The term "antisense" refers to the fact that such oligonucleotides are complementary to their intracellular targets, e.g., TRAG. (See for example, Jack Cohen, OLIGODEOXYNUCLEOTIDES, Antisense

Inhibitors of Gene Expression, CRC Press, 1989; and Synthesis 1:1-5 (1988)). The TRAG antisense oligonucleotides of the present invention include derivatives such as S-oligonucleotides (phosphorothioate derivatives or S-oligos, see, Jack Cohen, supra), which exhibit enhanced cancer cell growth inhibitory action.

S-oligos (nucleoside phosphorothioates) are isoelectronic analogs of an oligonucleotide (O-oligo) in which a nonbridging oxygen atom of the phosphate group is replaced by a sulfur atom. The S-oligos of the present invention may be prepared by treatment of the corresponding O-oligos with 3H-1,2-benzodithiol-3-one-1,1-dioxide, which is a sulfur transfer reagent. See Iyer, R. P. et al, J. Org. Chem. 55:4693-4698 (1990); and Iyer, R. P. et al., J. Am. Chem. Soc., 112:1253-1254 (1990), the disclosures of which are fully incorporated by reference herein.

The TRAG antisense oligonucleotides of the present invention may be RNA or DNA which is complementary to and stably hybridizes with the first 100 N-terminal codons or last 100 C-terminal codons of the TRAG genome or the corresponding mRNA. While absolute complementarity is not required, high degrees of complementarity are preferred. Use of an oligonucleotide complementary to this region allows for the selective hybridization to TRAG mRNA and not to mRNA specifying other regulatory subunits of protein kinase. Preferably, the TRAG antisense oligonucleotides of the present invention are a 15 to 30-mer fragment of the antisense DNA molecule having a sequence that hybridizes to TRAG mRNA. Optionally, TRAG antisense oligonucleotide is a 30-mer oligonucleotide that is complementary to a region in the first 10 N-terminal codons and last 10 C-terminal codons of TRAG. Alternatively, the antisense molecules are modified to employ ribozymes in the inhibition of TRAG expression. (L.A. Couture & D. T. Stinchcomb, Trends Genet., 12: 510–515 (1996)).

In one embodiment, the TRAG antisense oligonucleotide is coadministered with an agent that enhances the uptake of the antisense molecule by the cells. For example, the TRAG antisense oligonucleotide may be combined with a lipophilic cationic compound that may be in the form of liposomes. The use of liposomes to introduce nucleotides into cells is taught, for example, in U.S. Pat. Nos. 4,897,355 and 4,394,448, the disclosures of which are incorporated by reference in their entirety. See also U.S. Pat. Nos. 4,235,871, 4,231,877, 4,224,179, 4,753,788, 4,673,567, 4,247,411, and 4,814,270 for general

methods of preparing liposomes comprising biological materials. Alternatively, the TRAG antisense oligonucleotide may be combined with a lipophilic carrier such as any one of a number of sterols including cholesterol, cholate, and deoxycholic acid.

In another embodiment, the TRAG antisense oligonucleotide may be coadministered with a second agent that is affected by TRAG expression. In one embodiment, this second agent is one or more isoforms of TGF-β. In a preferred embodiment, a combination of TRAG antisense oligonucleotides and TGF-β1 are administered to cells that have reduced sensitivity to TGF-β due to TRAG overexpression. In this embodiment, a combination of these two molecules may be used to synergistically induce TGF-β1 mediated apoptosis. Methods pertaining to these embodiments are well known in the art. (Zwicker et al., Science, 271:1595-1597 (1996); Field et al., Cell, 85: 549-561 (1996); Slack et al., J. Cell Bio. ,129: 779-788 (1995); Hiebert et al., Mol Cell Biol., 15: 6864-6874 (1995); White, E., Genes Dev., 10: 1-15 (1996); Martin et al., Cell, 82:349-352 (1995)).

The antisense oligonucleotides of the present invention may be prepared according to any of the methods that are well known to those of ordinary skill in the art. Preferably, the antisense oligonucleotides are prepared by solid phase synthesis. (See Goodchild, J., Bioconjugate Chemistry, 1:165-167 (1990)), for a review of the chemical synthesis of oligonucleotides. Alternatively, the antisense oligonucleotides can be obtained from a number of companies that specialize in the custom synthesis of oligonucleotides.

3. Use of TRAG Nucleic Acids in the Generation of Transgenic Animals.

Nucleic acids that encode TRAG or its modified forms can also be used to generate either transgenic animals or "knockout" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal stage, e.g., an embryonic stage. A transgene is a DNA that is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding TRAG can be used to clone genomic DNA encoding TRAG in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells that

express DNA encoding TRAG. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for TRAG transgene incorporation with tissue–specific enhancers. Transgenic animals that include a copy of a transgene encoding TRAG introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding TRAG. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, if an animal is treated with the reagent and a reduced incidence of the pathological condition is observed as compared to untreated animals bearing the transgene, this would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, nonhuman homologues of TRAG can be used to construct a TRAG knockout animal that has a defective or altered gene encoding TRAG as a result of homologous recombination between the endogenous gene encoding TRAG and altered genomic DNA encoding TRAG introduced into an embryonic cell of the animal. For example, cDNA encoding TRAG can be used to clone genomic DNA encoding TRAG in accordance with established techniques. A portion of the genomic DNA encoding TRAG can be deleted or replaced with another gene, such as a gene encoding a selectable marker that can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector (see e.g., Thomas and Capecchi, Cell, 51:503 (1987) for a description of homologous recombination vectors). The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected (see e.g., Li et al., Cell, 69:915 (1992)). The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras (see e.g., Bradley, in Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113–152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a knockout animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance,

for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the TRAG polypeptide.

4. Use of TRAG Upstream Control Sequences For Evaluating Neoplastic Processes

The genomic TRAG control sequences of the present invention, whether positive, negative, or both, may be employed in numerous various combinations and organizations to assess the regulation of TRAG. Moreover, in the context of multiple unit embodiments and/or in embodiments that incorporate both positive and negative control units, there is no requirement that such units be arranged in an adjacent head-to-head or head-to-tail construction in that the improved regulation capability of such multiple units is conferred virtually independent of the location of such multiple sequences with respect to each other. Moreover, there is no requirement that each unit to comprise the same positive or negative element. All that is required is that such sequences be located upstream of and sufficiently proximal to a transcription initiation site.

To evaluate TRAG regulatory elements in the context of heterologous genes, one simply obtains the structural gene and locates one or more of such control sequences upstream of a transcription initiation site. Additionally, as is known in the art, it is generally desirable to include TATA-box sequences upstream of and proximal to a transcription initiation site of the heterologous structural gene. Such sequences may be synthesized and inserted in the same manner as the novel control sequences. Alternatively, one may desire to simply employ the TATA sequences normally associated with the heterologous gene. In any event, TATA sequences are most desirably located between about 20 and 30 nucleotides upstream of transcription initiation.

Preferably the heterologous gene is a reporter gene that encodes an enzyme that produces colorimetric or fluorometric change in the host cell, which is detectable by in situ analysis and which is a quantitative or semiquantitative function of transcriptional activation. Exemplary enzymes include esterases, phosphatases, proteases (tissue plasminogen activator or urokinase), and other enzymes capable of being detected by activity that generates a chromophore or fluorophore as will be known to those skilled in the art. A preferred example is E. coli beta-galactosidase. This enzyme produces a color change upon cleavage of the indigogenic substrate indolyl-B-D-galactoside by cells

bearing beta-galactosidase (*see, e.g.*, Goring et al., Science, 235:456-458 (1987) and Price et al., Proc. Natl. Acad. Sci. U.S.A., 84:156-160 (1987)). Thus, this enzyme facilitates automatic plate reader analysis of TRAG control sequence mediated expression directly in microtiter wells containing transformants treated with candidate activators. Also, because the endogenous beta-galactosidase activity in mammalian cells ordinarily is quite low, the analytic screening system using β-galactosidase is not hampered by host-cell background.

Another class of reporter genes that confer detectable characteristics on a host cell are those that encode polypeptides, generally enzymes, which render their transformants resistant against toxins, e.g., the neo gene, which protects host cells against toxic levels of the antibiotic G418; a gene encoding dihydrofolate reductase, which confers resistance to methotrexate; or the chloramphenicol acetyltransferase (CAT) gene (Osborne et al., Cell, 42:203-212 (1985)). Genes of this class are not preferred since the phenotype (resistance) does not provide a convenient or rapid quantitative output. Resistance to antibiotic or toxin requires days of culture to confirm or complex assay procedures if other than a biological determination is to be made.

5. Use of TRAG Polypeptides in Protein–Protein Interaction Studies.

Co-immunoprecipitation and Gal4 protein-protein interaction assays may be useful in screening for compounds modulating TRAG activity or in screening for compounds altering TRAG activity in a cell. For example, TRAG may participate with other WD repeat element-containing proteins in modulating TGF-β signaling. Those skilled in the art will understand that binding of a ligand at a molecular binding site can be modulated in a direct manner (e.g., by blocking the site), as well as altered in an indirect manner (e.g., by conformational changes induced following binding of a second (different) ligand at a distant site). In this regard, it is likely that the binding site specificity of TRAG for a particular WD repeat element-containing protein can be completely altered (i.e., to bind a different ligand) by agents that bind at distant sites in these proteins. A number of exemplary protocols that may be used in these studies are known in the art, see e.g. U.S. Patent No. 5,625,031.

Those skilled in the art will recognize that the functional regions of WD repeat element-containing proteins are particularly attractive targets for three-dimensional molecular modeling and for the construction of mimetic compounds, e.g., organic chemicals constructed to mimic the three-dimensional interactions between TRAG and another WD repeat element-containing protein. (See e.g., J. Wang, Curr Opin Gen. Dev, 7:39 (1997); Y. Taya et al., Trend Biochem. Sci., 22:14 (1997)).

6. Use of TRAG Containing Expression Vectors for Modulating Cellular Phenotype.

As discussed above, TRAG genes can be incorporated into any standard cloning vector. The term "vector" is well understood in the art and is synonymous with the often-used phrase "cloning vehicle." A suitable vector is a nonchromosomal double-stranded DNA comprising an intact replicon such that the vector is replicated when placed within a unicellular organism. Viral vectors include retroviruses, adenoviruses, adeno-associated virus (AAV), herpes virus, papovirus, adeno-retrovirus, etc. Other suitable vectors include plasmids. Plasmids and retroviruses are generally preferred as vectors.

As discussed in Example 8, pcDNA3.1-TRAG (pcDNA3.1 purchased from Invitrogen) was constructed and contained the CMV promoter. This promoter is suitable for expression of the TRAG genes in a wide variety of cells. However, the CMV promoter is not specifically required for transcription and expression of the TRAG genes. The CMV promoter can be replaced with other known promoters to improve the efficiency of transcription and expression in particular cells.

The promoter DNA can be amplified using PCR technology while concurrently providing restriction sites at the 5' and 3' ends of the promoter DNA. The amplified promoter DNA can then be inserted into a cloning vehicle (for example pcDNA3.1) using conventional endonucleases and known recombinant DNA technology. Cloning vectors containing the desired promoter upstream of the 5' end of TRAG genes may be constructed in this manner.

Cellular phenotype can be influenced by using the TRAG gene. In this context, cloning vectors containing an appropriate promoter and the TRAG gene may be constructed using PCR technology in a manner analogous to the preparation of vectors

containing exogenous genes as is well known in the art. Cloning vectors containing TRAG genes are transfected into host cells using known transfection processes. Suitable transfection processes include lipofection, electroporation, and retrovirus infection. When transfecting cells with TRAG, the desired cells are isolated and cultured in suitable media.

Transfection of cells using lipofection may be conducted according to standard lipofection procedures. (See Felgner et al., Proc. Natl. Acad. Sci. (USA), 84:7413-7417 (1987)). In general, liposome-mediated DNA transfection is accomplished by exposing 1-20 micrograms of plasmid DNA and commercially available liposomes (Bethesda Research Laboratories) in culture medium. The transfected cells are then repeatedly passaged in culture medium and the desired clones are isolated.

Retrovirus infection may also be accomplished using previously described procedures. (See, for example, Miller et al., J. Virol., 62:4337-4345 and Halbert et al., J. Virol., 65:473-378 (1991)). In general, plasmid DNA is transfected into a desired packaging cell line such as Psi-2 or other cell lines, using standard calcium phosphate precipitation. Viruses produced from the Psi-2 cells or equivalent cells are then used to infect an amphotropic packaging cell line, for example PA317. Viruses produced by the amphotropic packaging cell line are used to infect the desired host parent cells of the present invention.

Selection of clones with a modulated phenotype may be undertaken by a variety of protocols that are well known in the art (see e.g. U.S. Patent No. 5,376,542). In such selections, the cells may be selected for their ability to respond to factors such as TGF-β. Alternatively, cells can be selected by their ability to form colonies and grow in soft agar. Moreover, the cells can be selected by their ability to form tumors in animal models such as in nude mice. Similarly, cancer cells or cells having aggressive metastasis can be selected.

After transfection, the desired clones are selected by culturing in optimal media and repeated passaging. Generally, 10-20 passages are required to eliminate spurious cells and obtain pure clonal cells. Optimal media are selected according to the type of parent cell that is utilized. For lymphocytes, RPMI media is preferred; for fibroblasts,

DMEM media is preferred; and for epithelial cells, a serum-free medium such as keratinocyte growth medium (KGM) or SFM (Gibco Company) is preferred.

Selected colonies are then tested to verify the presence of TRAG DNA and the expression of TRAG genes. Verification is confirmed by standard Southern hybridization techniques and immunoprecipitation to determine the presence or quantity of expressed TRAG proteins.

7. Chromosomal Localization.

In Example 2, chromosomal localization of the human and murine TRAG genes is described. Chromosomal localization was done by FISH analysis (Stokke, T. et al., Genomics, 26: 134-137 (1995)). Murine TRAG maps to chromosome 18, and human TRAG maps to chromosome 18.

In other embodiments the invention provides diagnostic assays for determining chromosomal rearrangement of TRAG genes in a cell. The chromosomal location of TRAG genes is conveniently determined in chromosomal smears by *in situ* hybridization with oligonucleotide probes or cDNA and the like. Translocation of a TRAG gene, i.e., from a chromosomal location found in a normal cell to a location found in a transformed cell, may contribute to a phenotype of uncontrolled cell growth by removing normal transcription regulatory control of expression of TRAG. In the case where the rearrangement induces overexpression, the cell may acquire a malignant (i.e., uncontrolled) growth phenotype, and in the case where the rearrangement induces underexpression, the cell may undergo premature senescence. Screening cellular samples from individuals for the potential of TRAG chromosomal rearrangement may indicate a relative risk factor for the possibility of developing cancer.

E. Anti-TRAG Antibodies

The present invention further provides anti-TRAG antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

1. Polyclonal Antibodies

The TRAG antibodies may comprise polyclonal antibodies. Methods of

preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the TRAG polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants that may be employed include Freund's complete adjuvant and MPL–TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). Further, polyclonal antibodies may be generated commercially, for example by Genemed Synthesis, Inc. using art-accepted methods.

2. Monoclonal Antibodies

The TRAG antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the TRAG polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if nonhuman mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, *Monoclonal Antibodies: Principles and Practice*, Academic Press, (1986) pp. 59–103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine, and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin,

and thymidine ("HAT medium"), which are substances that can prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J., Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against TRAG. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods (Goding, supra). Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal. The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography. The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into

host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a nonimmunoglobulin polypeptide. Such a nonimmunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention or can be substituted for the variable domains of one antigen–combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking.

Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking. *In vitro* methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Humanized Antibodies

The TRAG antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of nonhuman (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains, or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$, or other antigen–binding subsequences of antibodies) that contain minimal sequence derived from nonhuman immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary–determining region (CDR) of the recipient are replaced by residues from a CDR of a nonhuman species (donor antibody) such as mouse, rat, or rabbit having the desired specificity, affinity, and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding nonhuman residues. Humanized antibodies may also comprise residues that are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized

antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a nonhuman immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., Nature, 321:522–525 (1986); Riechmann et al., Nature, 332:323–329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593–596 (1992)).

Methods for humanizing nonhuman antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source that is nonhuman. These nonhuman amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and coworkers (Jones et al., Nature, 321:522–525 (1986); Riechmann et al., Nature, 332:323–327 (1988); Verhoeyen et al., Science, 239:1534–1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a nonhuman species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86–95 (1991)).

Humanized antibodies can also be prepared according to the methods disclosed by, for example, U.S. Patent Nos. 5,175,384; 5,434,340; 5,545,806; 5,569,825; 5,591,669; 5,625,126; 5,633,425; 5,916,771; and 5,589,369.

4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one

of the binding specificities is for the TRAG, the other one is for any other antigen, and preferably for a cell–surface protein or receptor or receptor subunit. Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co–expression of two immunoglobulin heavy–chain/light–chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537–539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655–3659 (1991).

Antibody variable domains with the desired binding specificities (antibody–antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy–chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy–chain constant region (CH1) containing the site necessary for light–chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy–chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors and are cotransfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980) and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl–4–mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

F. Uses for TRAG Antibodies

The TRAG antibodies of the invention have various utilities. For example, TRAG antibodies may be used in diagnostic assays for TRAG, *e.g.*, detecting its expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays, and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases (Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147–158). The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^{3}H, ^{14}C, ^{32}P, ^{35}S, or ^{125}I, a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta–galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982). In addition, TRAG or antibodies that recognize TRAG may be used in drug screening assays to identify compounds that act to positively or negatively modulate the function of TRAG.

The antibodies can also be TRAG antagonists or agonists. Antibodies may also be useful therapeutically either alone, as agents that would act directly to interfere with the function of TRAG or indirectly as targeting agents capable of delivering a toxin, for example, pseudomonas exotoxin or radioisotopes, conjugated thereto to a desired site.

TRAG antibodies also are useful for the affinity purification of TRAG from recombinant cell culture or natural sources. In this process, the antibodies against TRAG are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the TRAG to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the TRAG, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the TRAG from the antibody.

****

The following examples are offered for illustrative purposes only and are not

intended to limit the scope of the present invention in any way. All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia, USA.

EXAMPLE 1

Isolation of Rat TRAG cDNA and Determination of Tissue Expression

Loss of TGF-β1 induced growth inhibition is an early event during spontaneous transformation of RLE cells. (A.C. Hugget et al., Cancer Res., 51:5929 (1991)). This resistance to the growth inhibitory effects of TGF-β1 can be caused by multiple factors. (E.R. Barrack, Prostate, 31:61 (1997); R.W. Padgett et al., Cytokine Growth Factor Rev., 8:1 (1997); L. Attisano, et al., Cytokine Growth Factor Rev., 7:327 (1996)). However, we observed a number of the transformed cell lines displaying resistance to the growth inhibitory effects of TGF-β1 that apparently had a "normal" number of TGF-β receptors. (A.C. Hugget et al., Cancer Res., 51:5929 (1991)). These observations suggested a post-receptor disruption of the growth inhibitory signal(s) of TGF-β1. To search for endogenous genes that confer resistance to the growth inhibitory effects of TGF-β1, and which in turn may lead to cellular transformation, subtractive hybridization (J.T. Woitach et al., Nat. Genet., 19:371 (1998)) was done on two rat liver epithelium (RLE)-derived cell lines that were sensitive (RLE phi 13) or insensitive (B5T) to TGF-β. (For a discussion about the production of RLE phi 13 and B5T cell lines, see A.C. Huggett et al., Cancer Research, 51:5929 (1991)). B5T designates a cell line derived from an RLE parent line by multiple passages (i.e. allowing cells to grow to fill a culture plate (confluency), and then diluting these cells into a fresh plate and allowing them to once again reach confluency). Unlike RLE phi 13 cells, B5T cells are insensitive to TGF-β and exhibit a transformed phenotype.

A novel transcript, TRAG (TGF-β Resistance Associated Gene), was identified and shown to be overexpressed in the B5T cells as well as one other transformed RLE line (C4T) derived in an identical way to B5T (A.C. Huggett et al., Cancer Research, 51:5929 (1991)), both of which are resistant to the effects of TGF-β1 (Figure 1).

Isolation of Rat TRAG cDNA

To isolate rat TRAG cDNA, poly-A^+ was first isolated from RLE phi 13 cells and B5T cells by using oligo (dt) cellulose, as previously described (Bradley JE, Bishop GA, St John T, Frelinger JA, Biotechniques, 6:114-116 (1988)). The cDNAs were synthesized using SuperScript reverse transcriptase (BRL) as recommended by the vendor; added with BstXI adaptor (InVitrogen); and used to produce RLE phi 13 and B5T cDNA libraries in pcDNAIneo.

For B5T cDNA enriched subtractive cDNA library construction, B5T and RLE phi 13 cDNAs were digested with HindIII+XbaI and BamHI+XhoI, respectively, and subjected to agarose gel purification. Digested RLE phi 13 cDNA fragments (20 μg) were digested further with AluI+RsaI and dephosphorylated, then hybridized with HindIII+XbaI digested B5T cDNA fragments (0.4 μg). The hybridization mixture was used to construct the B5T cDNA enriched subtractive cDNA library in BluescriptM13 with HindIII+XbaI protruding termini. The enriched library was plated and single colonies were isolated. The plasmid from each clone/colony was sequenced, and novel sequences were used to screen a panel of rat tumor samples and synchronized cells. A 0.8kb cDNA clone encoding part of the 3' untranslated region (3'UTR) of the TRAG gene was isolated. This clone was picked for its overexpression in B5T cells, relative to RLE cells, following screening by Northern hybridization techniques, and for its novel sequence. The single similar (homologous) sequence that was found in the sequence databases represented an uncharacterized cDNA clone from human brain, identified by the name KIAA0541 (GenBank Accession Number AB011113). Using primers designed against this sequence, and making use of reverse transcriptase (RT)-PCR and sequencing, the remainder of the 3' untranslated region and part of the coding region were isolated and characterized from rat (Table 1). RT-PCR describes a molecular biology technique by which isolated single-stranded complementary DNA (cDNA) is produced from total

mRNA or poly-A$^+$ RNA using the Reverse Transcriptase enzyme (SuperScript II from Gibco, BRL Life Technologies). This cDNA can then be used as a template for amplification in a PCR reaction. The template for RT-PCR was total RNA isolated from B5T cells (derived from RLE parent cells) or fresh brain tissue from rat.

The remainder of the coding region of rat TRAG, in particular, ~2kb of the 5' upstream region (including the transcription and translation start sites) was obtained using the RACE (random amplification of cDNA ends) technique, as described in the manufacturer's instructions accompanying the SMART® RACE kit (Clontech, Palo Alto, CA). The human TRAG sequence homologous to this region was *not* present in the KIAA0541 clone.

Sequence specific to rat brain TRAG cDNA was isolated as described above from cDNA obtained by RT (reverse transcriptase) experiments using total RNA extracted from Fisher F344 rat brain tissue (Table 1, bases 2876-2968).

The final size of the coding region for rat TRAG cDNA was 4,464 base pairs, encoding a protein of 1,488 amino acids.

Plasmids and PCR products were sequenced using the BigDye Terminator Ready Reaction Kit (cat# 4303149, Perkin-Elmer Applied Biosystems). The reactions were purified using Centriflac Gel Filtration Columns (cat# 42453, Edge Biosystems, Inc.). The samples were analyzed on a 377 ABI Prism DNA sequencer (Perkin-Elmer Applied Biosystems) to determine the primary nucleotide sequence. The predicted amino acid sequence having 1,488 amino acids (with a predicted Mr ~ 165,000 Daltons) was then derived using PCGENE. The derived amino acid sequence is indicated by the 3-letter code below the corresponding bases in Table 1.

Isolation of Mouse TRAG cDNA

The mouse TRAG gene was isolated by RT-PCR and RACE-PCR as described above for the rat TRAG gene, except that the primers used were designed against the rat TRAG sequence (which is closer to mouse TRAG that human is). Brain-specific sequence (Table 3, bases 2876-2971) was obtained as outlines above, using brain tissue from the SvJ129 mouse strain. Mouse genomic clones were isolated by screening a mouse BAC genomic libraries from the SvJ strain of mouse (performed by Genome

Systems, Inc.). The template for RT-PCR was total RNA isolated from fresh liver or brain for mouse. The template derived from mouse brain was used to sequence cDNA specific to brain tissue, which contained an additional coding region not found in the liver. The template derived from mouse liver was used to sequence cDNA specific to liver, which was shown to lack that coding region present in brain tissue.

The mouse cDNA sequence was determined as described above for rat cDNA. A cDNA clone having 4,467 base pairs was generated from mouse by RT-PCR and RACE-PCR (Table 3).

The predicted amino acid sequence having 1,489 amino acids (with a predicted Mr ~ 165,000 Daltons) was then derived using PCGENE. The derived amino acid sequence is indicated by the 3-letter code below the corresponding bases in Table 3.

Isolation of Human TRAG cDNA

The human TRAG gene was isolated from human poly-A^+ RNA obtained commercially from Clontech (cat# 6510-1, Palo Alto, CA) by RT-PCR, as described above, but using human-specific oligonucleotide primers which were obtained from sequence data deposited in GenBank for KIAA0541 (Accession Number AB011113). These primers were chosen because this GenBank sequence showed homology to the 3'UTR of the cDNA clone derived from the B5T cell described above for the isolation of the rat TRAG sequence. However, this human sequence (KIAA0541) was incomplete, lacking the 5' ~1.5 to ~2kb of the TRAG gene, including the transcription and translation start sites. The sequence of this missing region was determined by searching GenBank with the relevant mouse sequence (i.e. the first ~2kb of the mouse TRAG gene). This search yielded an uncharacterized genomic clone from chromosome 18 (containing both intronic and exonic sequence from the 5'-most region of the TRAG gene). By comparing mouse TRAG cDNA to this human genomic sequence, it was possible to detect the human exonic sequences and assemble them into a recognizable cDNA sequence. The entire human TRAG sequence was confirmed using RT-PCR of human poly-A^+ RNA (see above) and sequencing. The exon specific to brain, which was present in the KIAA0541 sequence, can be found in Table 2, bases 2924-3022.

A cDNA clone having 4,470 base pairs was generated from human by RT-PCR (Table 2). The predicted amino acid sequence having 1,490 amino acids (with a predicted Mr ~ 165,000 Daltons) was then derived using PCGENE. The derived amino acid sequence is indicated by the 3-letter code below the corresponding bases in Table 2.

The comparison of the cDNA sequence of TRAG from human, mouse, and rat indicates that TRAG is a highly conserved gene in all of these species, as is the amino acid sequence (Table 4).

Characterization of TRAG cDNA and Protein

DNA homology searches of GenBank revealed TRAG to be a unique transcript. Yet some homologous sequences were identified. A cDNA clone from human brain (KIAA0541, GenBank Accession No. AB011113) having approximately 87% nucleotide homology with the mouse TRAG gene was identified. A putative Drosophila G-protein subunit (GenBank Accession No. AL021086) having approximately 52% homology with the mouse TRAG protein was also identified using the BLAST search program. This homology is shown in Table 5. We compared the deduced amino acid sequence of TRAG with known proteins in the Swissprot database. TRAG contains no domains similar to other proteins. However, the deduced protein sequence does contain 4 WD repeat motifs (Tables 1, 2 and 3, underlined).

Based on the identification of the WD repeat motifs, TRAG may interact with the TGF−β pathway through these WD repeat motifs in a way similar to TRIP−1 and STRAP. Considering the means by which TRAG was isolated (i.e. by virtue of its overexpression in TGF−β resistant cells), it could play an important role in the TGF−β pathway and upregulation may interfere with normal TGF−β signaling. TRAG may interact directly with one of the TGF−β receptors and/or with another WD repeat−containing protein in the pathway, for example TRIP−1 and/or STRAP. We have already demonstrated an association of TRAG overexpression with a loss of TGF−β sensitivity and with an aggressive and malignant phenotype (see Examples below).

TRAG Expression

The expression of TRAG in three types of rat cells (B5T, RLE phi 13, and C4T) was analyzed by Northern blot (Figure 1A) to determine whether any difference in expression existed in the transformed (TGF-β resistant) cell lines relative to the nontransformed (TGF-β sensitive) RLE phi 13 cell line.

To determine TRAG expression levels in certain rat cells, RNA was first isolated from 3 rat cell lines (transformed B5T, transformed C4T, and untransformed RLE cells) (A.C. Huggett, et al., Cancer Res., 51:5929 (1991)), poly-A^+ mRNA selected using oligo (dt) cellulose (Bradley JE, Bishop GA, St John T, Frelinger JA (1988), Biotechniques, 6:114-116), and 5 μg (10 μg of total RNA) was then fractionated under denaturing conditions by electrophoresis on a 1% agarose gel, containing formaldehyde (0.22 M) and 1x MOPS (50 mM 3-[N-morpholino] phopanesulfonic acid and 1mM EDTA). Samples were electrophoresed for about 3 hours at 80V, blotted overnight onto nitrocellulose membranes, and cross-linked under ultraviolet light. Membranes were prehybridized in QuikHyb (Stratagene, La Jolla, CA) for 15 minutes at 68°C, labeled probe (0.8kb dsDNA against part of the 3'UTR; about $1x10^6$ cpm/ml) was added, and the membrane and probe were incubated for about 60 minutes at 68°C. The membrane was washed and then exposed to phosphoimager plates for 48 hours. The results are shown in Figure 1A.

The expressed TRAG levels were then quantified (Figure 1B). The results indicate that TGF-β resistant B5T cells express higher levels (up to 7 fold) of TRAG than do RLE phi 13 cells.

EXAMPLE 2

Chromosomal Localization of TRAG in Mouse and Human

Chromosomal localization of TRAG in mouse and human was done by FISH analysis (Stokke, T., Collins, C., Kuo, W., Kowbel, D., Shadrvan, F., Tanner, M., Kallionienmi, A., Kallioniemi, O., Pinkel, D., Deaven, L., and Gray, J ; A physical map of chromosome 20 established using fluorescent in situ hybridization (FISH) and digital image analysis, Genomics, 26: 134-137 (1995)). A FISH analysis generally involves probing chromosomes with a labeled DNA. The probes used to localize TRAG on

mouse and human chromosomes were prepared by labeling purified mouse and human DNA, respectively, with biotin (Random-Prime labeling kit: Boehringer-Mannheim Corp., Indianapolis, IN).

To localize TRAG on the mouse chromosome, a unique BAC clone (BAC21521) isolated from a 129SvJ mouse kidney library was used as a probe. This BAC clone was isolated by using a probe to the 3' untranslated region (UTR) of the rat TRAG gene (screening performed commercially by Genome Systems, Inc., St. Louis, MO).

To localize TRAG on the human chromosome, a ~5kb DNA probe complementary to the central portion of the human TRAG gene (exons 12-14, including intervening introns), obtained by PCR amplification of human genomic DNA, was used.

The biotin-labeled DNA probe was hybridized *in situ* to chromosomes derived from normal methotrexate-synthronized peripheral leukocyte cultures for human and to chromosomes derived from normal spleen cells in the case of mouse. Overall, forty-eight cells were studied.

The conditions used for hyridization, processing, analysis, direct flourescent signal localization, and banded chromosomes have been described previously in detail (Zimonjic et al., Cytogenet. Cell Genet., 65:184 (1994)).

The TRAG gene was localized to chromosome 18D.1-E.3 in the mouse (see Figure 2A) and to chromosome 18q21.1-22 in the human (see Figure 2B).

The long arm of human chromosome 18, designated 18q, and particularly the distal end of that region, has been found to be involved in a number of cancers and developmental disorders. The 18q region contains a number of genes which have been shown to play an integral role in various tumorigenic pathways by deletion studies. These genes include DCC (deleted in colorectal cancer) (L. Hendrick, et al., Genes Dev., 8:1174 (1994)), BCL-2 (Y Tsujimito et al., Science, 228:1440 (1985)), and DPC4 (deleted in pancreatic cancer) (SA Hahn et al., Science, 271:350 (1996)). TRAG sits between DCC and BCL-2 on chromosome 18q, a region that has also been shown to potentially be involved in a chromosomal rearrangement-associated mucosa-associated lymphoid tissue (MALT) lymphomas. (A Stoffel et al., Genes Chromosomes Cancer, 24:156 (1999)).

Inherited monosomy of 18q (i.e. loss of only one allele, so-called "18q$^-$ syndrome") has been associated with a variety of developmental and physiological defects, particularly those of the brain. (AD Kline et al., Am. J. Hum. Genet., 52:895

(1993)). Associated with 18q⁻ syndrome are various dysmorphic features as well as a string of neurological and other brain abnormalities, including microcephaly and abnormal MRI. (AD Kline et al., Am. J. Hum. Genet., 52:895 (1993)). The severity of defects appears to be correlated with the extent of the 18q deletion involved. For example, patients studied with deletions in the region of the TRAG gene showed more severe defects, particularly of brain development, than those carrying deletions that did not incorporate this area of the chromosome. (AD Kline et al., Am. J. Hum. Genet., 52:895 (1993)). This circumstantial evidence, coupled with the data showing the highest levels of TRAG expression in the brain (see Example 3), leads us to conclude that beyond a role in carcinogenesis, the TRAG gene may play a vital role in brain and mental development.

EXAMPLE 3

Determination of Expression of TRAG

To determine the normal expression of TRAG, a commercial multiple tissue RNA blot from rat (Clontech, Palo Alto, CA) was probed, utilizing Northern blotting techniques, with a radioactively labeled TRAG probe as described in Example 1.

The results are shown in Figure 3A. Figure 3A shows TRAG mRNA having a transcript size of 7.2 kb in spleen, lung, liver, muscle, and kidney, and TRAG mRNA having a transcript size of 3.5 kb in the testis. TRAG mRNA having a variety of transcript sizes (arrow heads in Figure 3A) was identified in brain suggesting alternative splicing.

Other TRAG expression tests were conducted using mouse and human dot blot membranes (Clontech, Palo Alto, CA), utililizing the identical Northern blot analysis protocol as detailed in Example 1, except that different probes were used. For human tissues, a probe of approximately 1.5 kb of human TRAG from the 5' end of the gene was used, and for mouse tissues, a probe of approximately 2.5 kb of mouse TRAG from the center of the gene was used.

Figure 3B shows the results for expression in human and Figure 3C shows the results for expression in mouse. Although the commercial multiple tissue RNA blot from rat did not show TRAG mRNA in heart (Figure 3A), the dot blot corresponding to expression in human did show mRNA for TRAG in heart (Figure 3B). The dot blot showed the most abundant TRAG mRNA expression in brain (Figures 3B and 3C). The dot blot also showed TRAG mRNA in all fetal tissues as well as all tissues during embryonal development (Figures 3B and 3C).

EXAMPLE 4

Analysis of TRAG Protein

To evaluate the expression of TRAG protein, protein extracts from three rat cell lines were studied (RLE phi 13, B5T, and C4T) by Western blot. Protein (40 μg) from each of the cell lines was collected by lysis of cells in RIPA lysis buffer (1% Tergitol NP-40, 0.5% sodium deoxycholate, 0.1% SDS, in 1X PBS). Inhibitors were added just before use (Complete, Mini; Boehringer-Mannheim, Burlingame, CA) according to manufacturer's instructions. Lysis buffer was added to plates after a brief wash with 1X PBS (containing inhibitors), cells were scraped from the plate and placed on ice for 5 minutes. After passing lysate though a 22 gauge syringe needle, samples were centrifuged at 4°C and 14,000g for 30 minutes. The supernatant was collected as crude extract and the protein concentration determined using the Bio-Rad Protein Assay (Bio-Rad, Hercules, CA). This protein (30μg to 40μg) was electrophoresed for 2.5h, 80V at room temperature on a polyacrylamide gel using a Novex XCell II Mini-cell System (Novex, San Diego, CA), according to manufacurers instructions, and then blotted onto nitrocellulose (25V, 2 hours). The nitrocellulose was incubated in a solution of 5% non-fat milk powder (NFMP), 3% BSA (bovine serum albumin), and 0.5% Tween-20 in Tris-buffered saline (TBS) for 1 hour at room temperature, washed 3 times in TBS + 0.05% Tween-20 (TBS-T), and placed in a solution of 5% NFMP, 3% BSA, and TBS-T (1°/2° Antibody Solution). To this was added the primary rabbit anti-TRAG antibody (1:1000 dilution; production of antibody described in Example 7) and the membrane incubated overnight at 4°C. After 3 TBS-T washes, secondary antibody (anti-rabbit horse radish peroxidase (HRP)-conjugated IgG; Santa Cruz Biotechnology, Santa Cruz, CA) was applied to the membrane in the same 1°/2° Antibody Solution used for the

primary antibody. Incubation was one hour at room temperature. After 3 TBS–T washes, the membrane was placed in ECL solution (Amersham, Arlington Heights, IL) for 1 minute, according to the supplied protocol, followed by exposure to Kodak X–Omat AR film for between 30 seconds and 2 minutes. The primary polyclonal anti–TRAG antibody that was used was raised in a rabbit against a peptide from the TRAG protein. This was done commercially by Genemed Biotechnologies Inc., San Francisco, CA, as described in Example 7.

The results are shown in Figure 4. Figure 4 shows TRAG protein expressed in all three cell lines. Yet B5T and C4T show greater amounts of protein than RLE phi 13. The TRAG protein appears to be about 100 kDa to about 105 kDa. Although this molecular weight is not immediately compatible with a coding sequence of approximately 4.4 kb (see Table 1), it is believed that alternative splicing produces a transcript of between about 3.3 kb and about 3.5 kb.

EXAMPLE 5

Determination of the Status of TRAG in Tumor Cell Lines

To determine the status of TRAG in tumor cell lines, protein extract and mRNA extract from a variety of tumor cell lines were studied by Western blot and Northern blot, respectively. In one study, protein extract and mRNA extract from cells that were chemically transformed using aflatoxin or virally transformed were examined. The chemically transformed cells were AFL–B8 and AFL–D8, and the virally transformed cells were 3611T2, 3611T5, and J2–14. These cells were produced in our laboratory according to JB McMahon et al., Cancer Res., 46:4665 (1986) and AC Huggett et al., Cancer Res., 51:5929 (1991). The levels of TRAG protein and mRNA in these cells were compared to levels of TRAG mRNA and protein in RLE phi 13 and B5T cells.

To determine TRAG protein level in these cells, a Western blot was conducted. Protein (40 μg) was isolated from each of the cell lines as described above in Example 4. The isolated protein was electrophoresed on a polyacrylamide gel and blotted onto nitrocellulose according to the methods described in Example 4. The nitrocellulose was then probed with the polyclonal antibody (1: 1,000 dilution) according to Example 4. The results are shown in Figure 5A. The level of TRAG protein correlated with the aggressiveness and metastatic ability of the particular cell lines. For example, J2–14, which is a highly aggressive metastatic tumor cell line transformed with c–*raf* and v–*myc*, shows the highest level of TRAG protein, while 3611T2, which is metastatic, shows

a level of TRAG protein that is greater than 3611T5, which is nonmetastatic. Aggressiveness and metastatic ability were determined by observation of tumors following injection into nude mice (PJ Worland, et al., Mol. Carcinog., 3:20–29 (1990)).

To determine TRAG mRNA level in these cells, a Northern blot was conducted as described above in Example 1. The results are shown in Figure 5A.

In another study, protein extract and mRNA extract from tumor cell lines derived from double transgenic c–*myc*/TGF–α mice (i.e., 221T1, 223B.3, 241T1, 241T3, and 263B.1) were evaluated. To determine TRAG protein level in these cells, a Western blot was conducted as described above. In this blot, the cancer cell lines were compared to B5T cells and to RLE phi 13 cells. The results are shown in Figure 5B. As for the transformed rat cell line, TRAG protein levels can be seen to be greatly elevated in these cells relative to RLE, and at comparable levels to B5T. This data strengthens the association between high TRAG levels and a transformed phenotype.

To determine TRAG mRNA level in these cells, a Northern blot was conducted as described above. The results are shown in Figure 5B.

In another study, protein extract and mRNA extract from the following human tumor cell lines were evaluated: Alex, Chang, FOCUS, HepG2, Huh–7, Sk–Hep–1, WRL–68, and HeLa. To determine the level of TRAG protein, a Western blot as described above was conducted. The results are shown in Figure 5C. Figure 5C shows a significant increase in TRAG protein in HepG2, which is an aggressive, transformed cell line.

To determine TRAG mRNA level in these cells, a Northern blot was conducted as described above. The results are shown in Figure 5C.

In yet another study, protein extract and mRNA extract from fresh, primary tumors taken from live double transgenic TGF–α/c–*myc* mice were evaluated and compared to the corresponding levels in RLE phi 13 and B5T cells. Figure 5D shows that liver tumors from double transgenic TGF–α/c–*myc* mice have elevated levels of TRAG protein relative to RLE phi 13. Approximately 75% of the primary liver tumors (T) isolated from TGF–α/c–*myc* mice showed higher levels of TRAG protein than did normal tissue (N) (Figure 5D).

EXAMPLE 6

The Determination of Cellular Localization of the TRAG Protein

The cellular localization of the TRAG protein was determined by confocal

microscopy. To make this determination, RLE phi 13 and B5T cells were transfected with empty (i.e. no TRAG insert) green fluorescent protein (GFP) vector (Clontech, Palo Alto, CA). These cells acted as the control. Other RLE phi 13 and B5T cells were transfected with the N-terminal GFP protein vector also containing the TRAG gene. This N-terminal fusion protein vector was constructed by inserting an EcoRI/SalI fragment encoding full-length TRAG into the pEGFP-N2 vector (Clontech, Palo Alto, CA) containing the green fluorescent protein (GFP). Transfection was carried out using LipofectAMINE PLUS reagent (Gibco, BRL Life Technologies, Rockville, MD), according to manufacturer's instructions. Microscopy conditions were: 20X zoom and 3% laser power (for B5T) or 10% laser power for RLE phi 13. The results are shown in Figures 6A and 6B.

Cells transfected with empty (pEGFP-N2) vector displayed GFP signal over the entire cell (Figures 6A, panels under "GFP-only"). But cells having the TRAG-GFP fusion protein vector showed only cytoplasmic fluorescence (Figure 6A, panels under "TRAG-GFP").

The cellular localization of the TRAG protein was confirmed by immunohistochemistry. Immunohistochemistry was conducted according to the standard procedures described in E. Harlow and D. Lane, "Using Antibodies: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1999). In particular, the TRAG protein in paraformaldehyde-fixed RLE phi 13 and B5T cells was probed using anti-TRAG polyclonal antibody (1:1,000 dilution). The polyclonal antibody was prepared according to Example 4. The results are shown in Figure 6B. Figure 6B shows that the TRAG protein localizes to the cytoplasm, but not to the nucleus. In fact, strongest staining is seen *around* the nucleus, within the cytoplasm (i.e. perinuclear).

These results were compared to staining for pre-immune (PI) serum as a negative control. PI serum describes the serum removed from the rabbit *before* it is immunized. That is, before the rabbit was challenged with the TRAG polypeptide and induced to make the anti-TRAG antibody. Figure 6B shows significantly lighter staining for pre-immune serum than for RLE phi 13 and B5T cells.

EXAMPLE 7

Preparation of Antibodies that Bind TRAG

Polyclonal antibodies to TRAG were generated commercially by Genemed Synthesis, Inc. using art-accepted methods. Initially, the rabbits were immunized with a polypeptide corresponding to amino acids 760–781 of the rat TRAG sequence (Table 1). Subsequent Western Blot experiments (Example 4) showed that the anti-TRAG antibody cross-reacted well with both mouse and human TRAG.

The following example illustrates preparation of monoclonal antibodies that can specifically bind TRAG.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, supra. Immunogens that may be employed include purified TRAG, fusion proteins containing TRAG, and cells expressing recombinant TRAG on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the TRAG immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1–100 micrograms. Alternatively, the immunogen is emulsified in MPL–TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect TRAG antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of TRAG. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against TRAG. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against TRAG is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-TRAG monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

EXAMPLE 8

Cloning of mouse TRAG gene in pcDNA3.1 +

The full length mouse TRAG gene was cloned into pcDNA3.1+ vector in three parts: a 5' EcoRI/Cla I fragment, a central ClaI/XhoI, and a 3' XhoI/XbaI fragment. In all cases, PCR-amplified fragments were cloned into pGEM®-T Easy vectors (Promega, Madison, WI) according to manufacturer's instructions and using E. coli XL-2 Blue cells supercompetent cells or methylation-deficienct SCS110 cells (in case of the methylation sensitive restriction enzymes XbaI or ClaI; both cell types available from Strategene, La Jolla, CA). Desired regions from this PCR fragment were then cut out of the vectors using the abovementioned site-specific restriction enzymes. These fragments were joined in consecutive steps by subcloning into a pGEM®-T Easy vector. Finally, the complete gene fragment was subcloned into pcDNA3.1+, under the CMV promoter. This provides a construct ready for transfection into any cell type for the purposes of determining the effect of overexpression of TRAG, for example, in RLE phi 13 cells, which normally have low levels of this protein (as demonstrated in Examples 1 and 4).

EXAMPLE 9

Generation of TRAG Transgenic and Knockout Mice

To elucidate more fully the potential role of TRAG in development and in carcinogenesis, the breeding of a knockout mouse is carried out. Part of the TRAG gene is deleted from a mouse line by standard knockout techniques, resulting in a truncated or nonexistent protein product (as described above in section D.3). If the complete removal of the TRAG gene product produces an embryonic lethal, a conditional knockout could be considered. This involves engineering the TRAG knockout construct in such a way

that the deletion of the critical region of the TRAG gene occurs only after embryogenesis. In this way, the role of TRAG in an adult animal can be examined.

To determine the role of TRAG in an aggressive phenotype in tumor cell lines, a transgenic mouse model can be studied. To study a transgenic mouse model, the TRAG gene is inserted into the mouse genome under the control of a strong, constitutively active promoter element (which causes large amounts of the protein to be made at all times) or a conditionally active promoter element (which would induce large amounts of TRAG only on a specific stimulus). (See, for example, C Jhappen et al., Cell, 61:1137 (1990)). TRAG overexpression is evaluated by Northern blot analysis, as described in Example 1. In particular, the transgenic mice are studied to determine if TRAG overexpression shows more rapid or enhanced tumor formation or more aggressive metastasis.

TABLES

Table 1 shows the cDNA (SEQ ID NO: 1) and amino acid (SEQ ID NO: 2) sequences of rat TRAG from liver as determined by double strand sequence analysis of PCR products amplified from RLE (rat liver epithelium) cell and Fisher F344 rat brain tissue cDNA. The WD repeat elements are indicated by the underline, and the two tyrosine phosphorylation motifs are indicated in boldface type.

```
                10        20        30        40        50
GAAGCTTTGACAGGTTTGAAAAACACAATGGCAGGAAACAGTCTAGTTCTGCCCATT
                                    MetAlaGlyAsnSerLeuValLeuProIle

 60        70        80        90        100       110
GTTCTCTGGGGCCGCAAAGCACCCACACACTGTATTTCGTCAATACTGCTGACAGAT
ValLeuTrpGlyArgLysAlaProThrHisCysIleSerSerIleLeuLeuThrAsp

   120       130       140       150       160       170
GATGGGGGCACGATTGTGACTGGATGCCACGATGGACAGATATGTCTCTGGGATCTC
AspGlyGlyThrIleValThrGlyCysHisAspGlyGlnIleCysLeuTrpAspLeu

       180       190       200       210       220
TCAGAAGAGCTGGAAGTTAATCCCCGAGCACTATTATTTGGTCACACAGCAGCCATC
SerGluGluLeuGluValAsnProArgAlaLeuLeuPheGlyHisThrAlaAlaIle

230       240       250       260       270       280
ACTTGTTTGTCAAAAGCCTGTGCTTCTGGAGACAAACAGTACACCGTGAGTGCGTCC
ThrCysLeuSerLysAlaCysAlaSerGlyAspLysGlnTyrThrValSerAlaSer

 290       300       310       320       330       340
GCAAATGGAGAGATGTGCCTCTGGGATGTGAATGACGGCAGATGTATTGAATTTACA
AlaAsnGlyGluMetCysLeuTrpAspValAsnAspGlyArgCysIleGluPheThr

     350       360       370       380       390
```

```
AAATTAGCCTGCACACATACTGGCATACAGTTCTACCAGTTCTCTGTTGGAAATCAG
LysLeuAlaCysThrHisThrGlyIleGlnPheTyrGlnPheSerValGlyAsnGln

400       410       420       430       440       450
CGAGAGGGAAGGCTTCTGTGCCATGGGCATTACCCTGAAATCCTCGTTGTGGATGCC
ArgGluGlyArgLeuLeuCysHisGlyHisTyrProGluIleLeuValValAspAla

 460       470       480       490       500       510
ACCAGCCTTGAGGTGTTGTATTCCTTGGTATCGAAGATCTCTCCAGACTGGATCAGC
ThrSerLeuGluValLeuTyrSerLeuValSerLysIleSerProAspTrpIleSer

    520       530       540       550       560       570
TCCATGAGYATYATCCGCTCTCACCGGACACAAGAGGACACTGTGGTAGCGCTGTCC
SerMetSerIleIleArgSerHisArgThrGlnGluAspThrValValAlaLeuSer

        580       590       600       610       620
GTGACAGGTATTCTGAAGGTGTGGATTGTGACCTCTGAAATTAGTGGATTGCAGGAC
ValThrGlyIleLeuLysValTrpIleValThrSerGluIleSerGlyLeuGlnAsp

630       640       650       660       670       680
ACTGAGCCAATATTTGAGGAGGAATCCAAACCAATTTATTGTCAGAATTGCCAAAGC
ThrGluProIlePheGluGluGluSerLysProIleTyrCysGlnAsnCysGlnSer

   690       700       710       720       730       740
CTCTCTTTTTGTGCATTCACACAGAGGTCGCTCTTGGTTGTATGCTCCAAGTACTGG
LeuSerPheCysAlaPheThrGlnArgSerLeuLeuValValCysSerLysTyrTrp

       750       760       770       780       790
AGGGTGTTCGATGCTGGCGACTACTCCCTGCTGTGTTCAGGTCCTAGTGAAGATGGA
ArgValPheAspAlaGlyAspTyrSerLeuLeuCysSerGlyProSerGluAspGly

800       810       820       830       840       850
CAGACATGGACTGGAGGGGACTTTGTGTCTGCAGACAAAGTCATTATTTGGACTGAA
GlnThrTrpThrGlyGlyAspPheValSerAlaAspLysValIleIleTrpThrGlu

  860       870       880       890       900       910
AACGGGCAGAGTTACATTTACAAACTCCCTGCCAGTTGCCTTCCAGCTAGTGATTCA
AsnGlyGlnSerTyrIleTyrLysLeuProAlaSerCysLeuProAlaSerAspSer

     920       930       940       950       960
TTCCGCAGTGACGTGGGGAAAGCAGTTGAAAATCTGATTCCTCCCGTGCAGCATAGC
PheArgSerAspValGlyLysAlaValGluAsnLeuIleProProValGlnHisSer

970       980       990      1000      1010      1020
CTCTTGGATCAGAAGGATAGAGAGTTGGTAATTTGTCCTCCTGTTACTCGGTTTTTC
LeuLeuAspGlnLysAspArgGluLeuValIleCysProProValThrArgPhePhe

1030      1040      1050      1060      1070      1080
TATGGATGCAAGGAATATTTGCATAAGCTACTAATTCAGGGTGATTCTTCTGGAAGG
TyrGlyCysLysGluTyrLeuHisLysLeuLeuIleGlnGlyAspSerSerGlyArg

   1090      1100      1110      1120      1130      1140
TTAAGTATTTGGAACATAGCAGACATAGCAGACAAACAGGAAGCCAATGAAGGGCTA
LeuSerIleTrpAsnIleAlaAspIleAlaAspLysGlnGluAlaAsnGluGlyLeu

      1150      1160      1170      1180      1190
AAAACGACAACTTGTATTAGTTTGCAAGATGCATTTGACAAACTGAAGCCTTGTCCT
LysThrThrThrCysIleSerLeuGlnAspAlaPheAspLysLeuLysProCysPro

1200      1210      1220      1230      1240      1250
GCTGGAATTATCGATCAGCTGAGTGTGATTCCAAACAGCAACGAACCACTTAAAGTA
AlaGlyIleIleAspGlnLeuSerValIleProAsnSerAsnGluProLeuLysVal
```

```
      1260       1270       1280       1290       1300       1310
ACTGCGAGTGTCTACATACCAGCACACGGGCGCCTTGTTTGCGGCCGGGAAGACGGA
ThrAlaSerValTyrIleProAlaHisGlyArgLeuValCysGlyArgGluAspGly

     1320       1330       1340       1350       1360
AGCATCATTATCGTCCCTGCCACCCAGACGGCCATAGTTCAGCTGCTGCAGGGAGAA
SerIleIleIleValProAlaThrGlnThrAlaIleValGlnLeuLeuGlnGlyGlu

1370       1380       1390       1400       1410       1420
CACATGCTCAGACGAGGTTGGCCCCCGCACAGAACCCTCCGTGGCCACCGGAACAAA
HisMetLeuArgArgGlyTrpProProHisArgThrLeuArgGlyHisArgAsnLys

 1430       1440       1450       1460       1470       1480
GTCACGTGTTTGCTGTATCCTCATCAGGTCTCAGCTCGGTATGATCAAAGATACCTG
ValThrCysLeuLeuTyrProHisGlnValSerAlaArgTyrAspGlnArgTyrLeu

     1490       1500       1510       1520       1530
ATATCCGGAGGTGTGGATTTTTCCGTCATCATTTGGGACATTTTTTCCGGAGAAATG
IleSerGlyGlyValAspPheSerValIleIleTrpAspIlePheSerGlyGluMet

1540       1550       1560       1570       1580       1590
AAACATATCTTCTGTGTTCATGGTGGTGAGATCACACAACTTCTGGTCCCGCCAGAA
LysHisIlePheCysValHisGlyGlyGluIleThrGlnLeuLeuValProProGlu

1600       1610       1620       1630       1640       1650
AACTGTAGTGCAAGAGTTCAACACTGCGTCTGTTCTGTGGCCAGTGACCACTCTGTA
AsnCysSerAlaArgValGlnHisCysValCysSerValAlaSerAspHisSerVal

   1660       1670       1680       1690       1700       1710
GGGCTGCTAAGTCTGCGAGAGAAAAAATGCATCATGCTGGCGTCTCGTCACCTGTTC
GlyLeuLeuSerLeuArgGluLysLysCysIleMetLeuAlaSerArgHisLeuPhe

      1720       1730       1740       1750       1760
CCTATTCAGGTGATCAAGTGGAGGCCTTCTGACGACTACCTGGTGGTGGGGTGCACG
ProIleGlnValIleLysTrpArgProSerAspAspTyrLeuValValGlyCysThr

1770       1780       1790       1800       1810       1820
GACGGCTCTGTGTGTGTCTGGCAGATGGACACTGGTGCGCTGGACCGCTGTGCAATG
AspGlySerValCysValTrpGlnMetAspThrGlyAlaLeuAspArgCysAlaMet

  1830       1840       1850       1860       1870       1880
GGGATAACAGCCGTGGAGATTCTCAATGCTTGTGACGAAGCTGTCCCTGCAGCAGTG
GlyIleThrAlaValGluIleLeuAsnAlaCysAspGluAlaValProAlaAlaVal

     1890       1900       1910       1920       1930
GACTCACTCAGTCACCCAGCAGTCAACCTGAAGCAAGCCATGACACGGCGGAGTCTC
AspSerLeuSerHisProAlaValAsnLeuLysGlnAlaMetThrArgArgSerLeu

1940       1950       1960       1970       1980       1990
GCCGCCCTTAAAAACATGGCCCACCACAAGCTGCAAACCCTTGCAACTAACCTTTTG
AlaAlaLeuLysAsnMetAlaHisHisLysLeuGlnThrLeuAlaThrAsnLeuLeu

 2000       2010       2020       2030       2040       2050
GCTTCTGAGGCCTCTGACAAGGGGAATTTACCTAAATATTCTCATAACTCCCTGATG
AlaSerGluAlaSerAspLysGlyAsnLeuProLysTyrSerHisAsnSerLeuMet

    2060       2070       2080       2090       2100
GTTCAAGCAATAAAGACAAACCTAACTGACCCGGATATCCATGTGCTTTTCTTTGAT
ValGlnAlaIleLysThrAsnLeuThrAspProAspIleHisValLeuPhePheAsp

2110       2120       2130       2140       2150       2160
GTGGAAGCTTTGATTATTCAACTCCTGACTGAAGAAGCCTCTAGGCCGAATACTGCA
ValGluAlaLeuIleIleGlnLeuLeuThrGluGluAlaSerArgProAsnThrAla
```

```
                2170       2180       2190       2200       2210       2220
                CTTATTTCCCCAGAGAATCTGCAGAAAGCATCTGGCAGTTCAGACAAAGGGGGCTCT
                LeuIleSerProGluAsnLeuGlnLysAlaSerGlySerSerAspLysGlyGlySer

                   2230       2240       2250       2260       2270       2280
                TTCCTGACTGGAAAACGAGCGGCAGTTCTTTTCCAGCAAGTGAAAGAAACTATCAAA
                PheLeuThrGlyLysArgAlaAlaValLeuPheGlnGlnValLysGluThrIleLys

                      2290       2300       2310       2320       2330
                GAGAACATAAAGGAGCACCTCCTTGATGAGGAGGAGGACGAGGAAGAGGTGATGAGG
                GluAsnIleLysGluHisLeuLeuAspGluGluGluAspGluGluGluValMetArg

             2340       2350       2360       2370       2380       2390
                CAGAGGAGGGAAGAAAGTGACCCTGAGTACCGGGCCAGCAAGTCCAAACCACTTACC
                GlnArgArgGluGluSerAspProGluTyrArgAlaSerLysSerLysProLeuThr

                  2400       2410       2420       2430       2440       2450
                CTACTAGAATACAACCTTACTATGGATACCGCAAAATTATTCATGTCCTGTCTCCAC
                LeuLeuGluTyrAsnLeuThrMetAspThrAlaLysLeuPheMetSerCysLeuHis

                     2460       2470       2480       2490       2500
                GCCTGGGGTTTGAATGAAGTTCTGGATGAAGTTTGCCTCGATCGCCTCGGCATGCTG
                AlaTrpGlyLeuAsnGluValLeuAspGluValCysLeuAspArgLeuGlyMetLeu

           2510       2520       2530       2540       2550       2560
                AAACCACACTGCACAGTGTCCTTTGGTCTCCTATCCAGAGGAGGTCATATGTCCTTG
                LysProHisCysThrValSerPheGlyLeuLeuSerArgGlyGlyHisMetSerLeu

                  2570       2580       2590       2600       2610       2620
                ATGCTTCCTGGTTATAATCAGGCTGCTGGAAAGCTACTGCAGGCGAAAGCAGAAGCA
                MetLeuProGlyTyrAsnGlnAlaAlaGlyLysLeuLeuGlnAlaLysAlaGluAla

                     2630       2640       2650       2660       2670
                GGACGGAAGGGGCCAGCAACGGAGAGCGTAGGCAAGGGGACCTACACAGTGTCCCGA
                GlyArgLysGlyProAlaThrGluSerValGlyLysGlyThrTyrThrValSerArg

           2680       2690       2700       2710       2720       2730
                GCGGTCACCACGCAACATCTGTTGTCCATCATATCTCTGGCAAATACTTTAATGAGC
                AlaValThrThrGlnHisLeuLeuSerIleIleSerLeuAlaAsnThrLeuMetSer

                2740       2750       2760       2770       2780       2790
                ATGACCAATGCGACGTTCATTGGAGATCACATGAAGAAGGGCCCCACCAGGCCGCCT
                MetThrAsnAlaThrPheIleGlyAspHisMetLysLysGlyProThrArgProPro

                   2800       2810       2820       2830       2840       2850
                AGACCAGGCACCCCAGACCTTTCTAAGGCAAGGGATTCCCCCCCCAGCCTCCAGTAAC
                ArgProGlyThrProAspLeuSerLysAlaArgAspSerProProAlaSerSerAsn

                      2860       2870       2880       2890       2900
                ATTGTGCAAGGACAGATTAAACAAGCTGCTGCGCCTGTCTCTGCTCGGTCTGCCGCC
                IleValGlnGlyGlnIleLysGlnAlaAlaAlaProValSerAlaArgSerAlaAla

             2910       2920       2930       2940       2950       2960
                GACCACTCTGGCTCTGCCTCTGCCTCTCCTGCTTTACGTACCTGCTTCTTAGTGAAT
                AspHisSerGlySerAlaSerAlaSerProAlaLeuArgThrCysPheLeuValAsn

                  2970       2980       2990       3000       3010       3020
                GAAGGATGGAGCCAACTAGCTGCCATGCACTGTGTCATGCTGCCGGACCTGCTGGGG
                GluGlyTrpSerGlnLeuAlaAlaMetHisCysValMetLeuProAspLeuLeuGly

                     3030       3040       3050       3060       3070
                CTGGGTAAATTCAGGCCTCCTCTTCTGGAGATGCTAGCTCGAAGATGGCAAGATCGG
```

```
                LeuGlyLysPheArgProProLeuLeuGluMetLeuAlaArgArgTrpGlnAspArg

              3080       3090       3100       3110       3120       3130
                TGCTTGGAGGTGAGAGAGGCTGCACAGGCCCTTCTTCTAGCAGAGCTGAGAAGAATT
                CysLeuGluValArgGluAlaAlaGlnAlaLeuLeuLeuAlaGluLeuArgArgIle

                 3140       3150       3160       3170       3180       3190
                GAGCAGGCAGGACGGAAGGAGACTATTGATACCTGGGCTCCTTACTTACCTCAGTAC
                GluGlnAlaGlyArgLysGluThrIleAspThrTrpAlaProTyrLeuProGlnTyr

                    3200       3210       3220       3230       3240
                ATGGACCATGTCATATCACCTGGAGTCACGGCGGAAGCCATGCAGACTATGGCAGCT
                MetAspHisValIleSerProGlyValThrAlaGluAlaMetGlnThrMetAlaAla

            3250       3260       3270       3280       3290       3300
                GCTCCAGATGCCTCGGGGCCAGAAGCCAAAGTCCAGGAAGAAGAGCATGACCTCGTG
                AlaProAspAlaSerGlyProGluAlaLysValGlnGluGluGluHisAspLeuVal

                3310       3320       3330       3340       3350       3360
                GACGATGACATCACCACTGGTTGCTTATCAAGTGTCCCACAAATGAAAAAGATGTCC
                AspAspAspIleThrThrGlyCysLeuSerSerValProGlnMetLysLysMetSer

                   3370       3380       3390       3400       3410       3420
                ACATCTTACGAAGAAAGAAGGAAGCAGGCCACTGCTATTGTTCTCCTGGGAGTGATA
                ThrSerTyrGluGluArgArgLysGlnAlaThrAlaIleValLeuLeuGlyValIle

                      3430       3440       3450       3460       3470
                GGAGCAGAGTTTGGAGCTGAAATTGAACCACCAAAACTGCTGACCAGACCTCGGAGC
                GlyAlaGluPheGlyAlaGluIleGluProProLysLeuLeuThrArgProArgSer

              3480       3490       3500       3510       3520       3530
                TCTAGTCAAATTCCTGAAGGATTTGGTTTGACAAGTGGAGGTTCCAACTACTCTCTG
                SerSerGlnIleProGluGlyPheGlyLeuThrSerGlyGlySerAsnTyrSerLeu

                 3540       3550       3560       3570       3580       3590
                GCCAGACATACGTGCAAGGCACTGACATTTCTTCTGCTACAGCCACCAAGTCCCAAA
                AlaArgHisThrCysLysAlaLeuThrPheLeuLeuLeuGlnProProSerProLys

                     3600       3610       3620       3630       3640
                CTTCCTCCTCATAGCACCATCCGGAGAACTGCCATTGACCTGATCGGGCGAGGGTTC
                LeuProProHisSerThrIleArgArgThrAlaIleAspLeuIleGlyArgGlyPhe

              3650       3660       3670       3680       3690       3700
                ACCGTGTGGGAGCCTTACATGGACGTGTCTGCTGTCCTGATGGGGCTGCTGGAGCTG
                ThrValTrpGluProTyrMetAspValSerAlaValLeuMetGlyLeuLeuGluLeu

                 3710       3720       3730       3740       3750       3760
                TGTGCAGATGCTGAGAAGCAGCTGGCCAACATCACAATGGGGCTGCCTCTGAGCCCT
                CysAlaAspAlaGluLysGlnLeuAlaAsnIleThrMetGlyLeuProLeuSerPro

                    3770       3780       3790       3800       3810
                GCAGCTGACTCTGCCCGATCCGCAAGACACGCCCTTTCTCTCATAGCCACCGCCAGA
                AlaAlaAspSerAlaArgSerAlaArgHisAlaLeuSerLeuIleAlaThrAlaArg

            3820       3830       3840       3850       3860       3870
                CCACCCGCCTTCATCACCACCATAGCTAAGGAGGTGCACAGACACACGGCCCTTGCA
                ProProAlaPheIleThrThrIleAlaLysGluValHisArgHisThrAlaLeuAla

                3880       3890       3900       3910       3920       3930
                GCAAATACCCAGTCCCAGCAGAGTATCCACACCACCACACTGGCAAGGGCTAAAGGC
                AlaAsnThrGlnSerGlnGlnSerIleHisThrThrThrLeuAlaArgAlaLysGly

                   3940       3950       3960       3970       3980       3990
```

```
GAAATCCTGAGAGTCATTGAAATTCTTATCGAAAAGATGCCTACGGATGTTGTGGAT
GluIleLeuArgValIleGluIleLeuIleGluLysMetProThrAspValValAsp

      4000      4010      4020      4030      4040
CTTCTTGTGGAGGTCATGGACATCATCATGTACTGCCTGGAAGGATCTTTAGTTAAG
LeuLeuValGluValMetAspIleIleMetTyrCysLeuGluGlySerLeuValLys

4050      4060      4070      4080      4090      4100
AAGAAGGGTCTTCAGGAGTGTTTCCCAGCCATCTGCAGGTTCTACATGGTCAGCTAT
LysLysGlyLeuGlnGluCysPheProAlaIleCysArgPheTyrMetValSerTyr

  4110      4120      4130      4140      4150      4160
TATGAGCGGAGTCACAGAATTGCAGTTGGAGCACGCCATGGCTCAGTGGCCCTGTAT
TyrGluArgSerHisArgIleAlaValGlyAlaArgHisGlySerValAlaLeuTyr

     4170      4180      4190      4200      4210
GACATCCGGACTGGGAAATGTCAGACAATCCACGGACACAAGGGACCTATCACTGCA
AspIleArgThrGlyLysCysGlnThrIleHisGlyHisLysGlyProIleThrAla

4220      4230      4240      4250      4260      4270
GTGTCCTTTGCTCCTGATGGGCGTTACCTTGCCACCTACTCAAACACTGACAGCCAC
ValSerPheAlaProAspGlyArgTyrLeuAlaThrTyrSerAsnThrAspSerHis

 4280      4290      4300      4310      4320      4330
ATTTCTTTCTGGCAGATGAACACCTCACTTCTGGGAAGCATTGGCATGCTGAACTCA
IleSerPheTrpGlnMetAsnThrSerLeuLeuGlySerIleGlyMetLeuAsnSer

    4340      4350      4360      4370      4380
GCACCTCAGCTGCGCTGCATCAAGACCTACCAGGTACCTCCAGTGCAGCCCGCATCC
AlaProGlnLeuArgCysIleLysThrTyrGlnValProProValGlnProAlaSer

4390      4400      4410      4420      4430      4440
CCTGGCTCGCACAACGCCCTCAGGTTGGCCCGGCTCATCTGGACTTCCAACCGGAAT
ProGlySerHisAsnAlaLeuArgLeuAlaArgLeuIleTrpThrSerAsnArgAsn

4450      4460      4470      4480      4490
GTTATCCTCATGGCCCACGATGGGAAGGAGCACCGCTTCATGGTCTGA
ValIleLeuMetAlaHisAspGlyLysGluHisArgPheMetValSTP
```

Table 2 shows the cDNA (SEQ ID NO: 3) and amino acid (SEQ ID NO: 4) sequences of human TRAG from brain as determined by double strand sequence of PCR products amplified from human liver tissue cDNA. The WD repeat elements are indicated by the underline, and the two tyrosine phosphorylation motifs are indicated in boldface type.

```
        10        20        30        40        50
TAACATGTTTTCGAGAGAAATATTGTAATATCTGACAATTTTTATAACATTTTCAGG

 60        70        80        90       100       110
TTTGAAAACACAAACACAATGGCAGGAAACAGCCTTGTTCTACCCATTGTTCTTTGG
                  MetAlaGlyAsnSerLeuValLeuProIleValLeuTrp

   120       130       140       150       160       170
GGTCGAAAAGCGCCCACACATTGCATCTCAGCCGTACTTTTAACAGATGATGGGGCC
GlyArgLysAlaProThrHisCysIleSerAlaValLeuLeuThrAspAspGlyAla

      180       190       200       210       220
ACGATCGTAACAGGATGTCACGACGGACAAATATGTCTCTGGGATCTTTCAGTAGAA
ThrIleValThrGlyCysHisAspGlyGlnIleCysLeuTrpAspLeuSerValGlu
```

```
230        240        250        260        270        280
CTGCAAGTTAATCCTCGAGCACTGTTGTTTGGTCATACAGCATCAATCACTTGTTTG
LeuGlnValAsnProArgAlaLeuLeuPheGlyHisThrAlaSerIleThrCysLeu

  290        300        310        320        330        340
TCTAAAGCTTGTGCTTCCAGTGACAAACAGTATATTGTGAGTGCATCTGAAAGTGGA
SerLysAlaCysAlaSerSerAspLysGlnTyrIleValSerAlaSerGluSerGly

      350        360        370        380        390
GAGATGTGCCTCTGGGATGTGAGTGATGGCAGATGTATTGAATTTACAAAATTAGCT
GluMetCysLeuTrpAspValSerAspGlyArgCysIleGluPheThrLysLeuAla

400        410        420        430        440        450
TGCACACATACTGGCATACAGTTCTACCAGTTCTCTGTTGGGAATCAGCGAGAAGGA
CysThrHisThrGlyIleGlnPheTyrGlnPheSerValGlyAsnGlnArgGluGly

 460        470        480        490        500        510
AGGCTTTTATGCCACGGACATTACCCTGAAATCCTTGTTGTGGATGCTACCAGCCTT
ArgLeuLeuCysHisGlyHisTyrProGluIleLeuValValAspAlaThrSerLeu

    520        530        540        550        560        570
GAAGTATTATACTCCTTAGTATCAAAGATATCACCAGACTGGATTAGCTCCATGAGT
GluValLeuTyrSerLeuValSerLysIleSerProAspTrpIleSerSerMetSer

        580        590        600        610        620
ATTATTCGATCCCACCGAACACAAGAGGACACAGTGGTAGCACTCTCGGTGACTGGC
IleIleArgSerHisArgThrGlnGluAspThrValValAlaLeuSerValThrGly

630        640        650        660        670        680
ATCCTGAAGGTCTGGATTGTTACCTCGGAAATAAGTGACATGCAGGATACTGAGCCA
IleLeuLysValTrpIleValThrSerGluIleSerAspMetGlnAspThrGluPro

   690        700        710        720        730        740
ATATTTGAGGAGGAATCCAAACCAATTTATTGTCAGAATTGCCAAAGCATCTCTTTT
IlePheGluGluGluSerLysProIleTyrCysGlnAsnCysGlnSerIleSerPhe

       750        760        770        780        790
TGTGCATTTACACAAAGGTCACTTTTGGTTGTGTGTTCCAAATATTGGAGGGTGTTC
CysAlaPheThrGlnArgSerLeuLeuValValCysSerLysTyrTrpArgValPhe

800        810        820        830        840        850
GATGCCGGAGACTATTCCTTGTTGTGTTCAGGTCCTAGTGAAAATGGACAGACATGG
AspAlaGlyAspTyrSerLeuLeuCysSerGlyProSerGluAsnGlyGlnThrTrp

  860        870        880        890        900        910
ACCGGGGGGGACTTTGTCTCATCAGATAAAGTCATCATTTGGACAGAAAATGGGCAA
ThrGlyGlyAspPheValSerSerAspLysValIleIleTrpThrGluAsnGlyGln

     920        930        940        950        960
AGTTATATTTACAAACTACCTGCCAGTTGCCTTCCAGCTAGTGATTCATTCCGCAGT
SerTyrIleTyrLysLeuProAlaSerCysLeuProAlaSerAspSerPheArgSer

970        980        990       1000       1010       1020
GATGTGGGGAAGGCAGTTGAAAATTTAATTCCTCCTGTACAACATATCCTCTTGGAT
AspValGlyLysAlaValGluAsnLeuIleProProValGlnHisIleLeuLeuAsp

1030       1040       1050       1060       1070       1080
CGAAAAGATAAAGAGTTGCTAATTTGTCCTCCTGTTACTCGGTTCTTCTATGGATGC
ArgLysAspLysGluLeuLeuIleCysProProValThrArgPhePheTyrGlyCys

   1090       1100       1110       1120       1130       1140
AGAGAATATTTCCATAAACTGTTAATTCAGGGTGATTCTTCTGGAAGGTTGAATATT
```

```
ArgGluTyrPheHisLysLeuLeuIleGlnGlyAspSerSerGlyArgLeuAsnIle

      1150      1160      1170      1180      1190
TGGAACATATCAGACACAGCTGATAAACAGGGAAGTGAAGAAGGGCTGGCAATGACA
TrpAsnIleSerAspThrAlaAspLysGlnGlySerGluGluGlyLeuAlaMetThr

1200      1210      1220      1230      1240      1250
ACTTCTATTAGTTTGCAAGAGGCATTTGATAAACTGAATCCTTGTCCTGCTGGAATT
ThrSerIleSerLeuGlnGluAlaPheAspLysLeuAsnProCysProAlaGlyIle

  1260      1270      1280      1290      1300      1310
ATAGATCAGCTGAGTGTGATTCCCAATAGTAATGAACCTCTTAAAGTAACTGCAAGT
IleAspGlnLeuSerValIleProAsnSerAsnGluProLeuLysValThrAlaSer

      1320      1330      1340      1350      1360
GTGTACATACCAGCACATGGACGACTTGTTTGTGGTCGTGAAGATGGAAGCATAGTT
ValTyrIleProAlaHisGlyArgLeuValCysGlyArgGluAspGlySerIleVal

1370      1380      1390      1400      1410      1420
ATTGTACCTGCCACACAGACGGCCATAGTACAGCTGTTGCAAGGGGAACACATGCTC
IleValProAlaThrGlnThrAlaIleValGlnLeuLeuGlnGlyGluHisMetLeu

 1430      1440      1450      1460      1470      1480
AGAAGAGGTTGGCCACCTCACAGAACACTCCGTGGTCATCGGAACAAAGTCACATGT
ArgArgGlyTrpProProHisArgThrLeuArgGlyHisArgAsnLysValThrCys

     1490      1500      1510      1520      1530
TTGCTATATCCTCATCAGGTCTCAGCTCGGTATGATCAAAGATACCTGATATCTGGA
LeuLeuTyrProHisGlnValSerAlaArgTyrAspGlnArgTyrLeuIleSerGly

1540      1550      1560      1570      1580      1590
GGTGTGGATTTTTCAGTCATAATTTGGGACATATTTTCTGGAGAAATGAAACATATC
GlyValAspPheSerValIleIleTrpAspIlePheSerGlyGluMetLysHisIle

1600      1610      1620      1630      1640      1650
TTCTGTGTTCATGGTGGTGAGATTACTCAACTTCTAGTTCCACCTGAAAACTGTAGT
PheCysValHisGlyGlyGluIleThrGlnLeuLeuValProProGluAsnCysSer

   1660      1670      1680      1690      1700      1710
GCAAGAGTACAGCACTGCATCTGCTCTGTAGCCAGTGACCACTCAGTAGGACTTCTA
AlaArgValGlnHisCysIleCysSerValAlaSerAspHisSerValGlyLeuLeu

      1720      1730      1740      1750      1760
AGTTTGCGAGAGAAAAAATGCATAATGTTGGCATCTCGTCACCTTTTTCCTATTCAA
SerLeuArgGluLysLysCysIleMetLeuAlaSerArgHisLeuPheProIleGln

1770      1780      1790      1800      1810      1820
GTAATCAAATGGAGGCCTTCTGATGATTACCTGGTGGTGGGGTGTTCAGATGGTTCT
ValIleLysTrpArgProSerAspAspTyrLeuValValGlyCysSerAspGlySer

  1830      1840      1850      1860      1870      1880
GTGTACGTCTGGCAAATGGATACTGGTGCATTGGATCGTTGTGTGATGGGGATAACA
ValTyrValTrpGlnMetAspThrGlyAlaLeuAspArgCysValMetGlyIleThr

      1890      1900      1910      1920      1930
GCAGTTGAGATTCTAAACGCTTGTGATGAAGCTGTTCCTGCTGCTGTTGATTCACTT
AlaValGluIleLeuAsnAlaCysAspGluAlaValProAlaAlaValAspSerLeu

1940      1950      1960      1970      1980      1990
AGTCATCCAGCAGTCAACCTAAAACAAGCTATGACGAGACGTAGTCTTGCTGCTCTT
SerHisProAlaValAsnLeuLysGlnAlaMetThrArgArgSerLeuAlaAlaLeu

 2000      2010      2020      2030      2040      2050
```

```
AAAAATATGGCCCATCATAAGCTACAAACCCTTGCAACTAACCTCTTGGCTTCTGAG
LysAsnMetAlaHisHisLysLeuGlnThrLeuAlaThrAsnLeuLeuAlaSerGlu

    2060       2070       2080       2090       2100
GCATCTGACAAGGGAAATTTACCTAAATATTCTCATAACTCCCTGATGGTTCAAGCA
AlaSerAspLysGlyAsnLeuProLysTyrSerHisAsnSerLeuMetValGlnAla

2110       2120       2130       2140       2150       2160
ATAAAGACAAACCTAACAGACCCGGACATACATGTGCTATTCTTTGATGTGGAAGCG
IleLysThrAsnLeuThrAspProAspIleHisValLeuPhePheAspValGluAla

2170       2180       2190       2200       2210       2220
TTGATTATTCAACTCCTGACTGAAGAAGCCTCTAGGCCGAATACTGCTCTTATTTCC
LeuIleIleGlnLeuLeuThrGluGluAlaSerArgProAsnThrAlaLeuIleSer

   2230       2240       2250       2260       2270       2280
CCAGAGAATTTGCAAAAAGCATCTGGCAGTTCAGACAAAGGGGGCTCTTTTTTAACT
ProGluAsnLeuGlnLysAlaSerGlySerSerAspLysGlyGlySerPheLeuThr

      2290       2300       2310       2320       2330
GGAAAACGAGCAGCAGTTCTCTTCCAACAAGTGAAAGAAACGATCAAAGAGAACATC
GlyLysArgAlaAlaValLeuPheGlnGlnValLysGluThrIleLysGluAsnIle

2340       2350       2360       2370       2380       2390
AAGGAACACCTCCTTGATGATGAAGAGGAGGATGAGGAGATAATGAGGCAGAGAAGG
LysGluHisLeuLeuAspAspGluGluGluAspGluGluIleMetArgGlnArgArg

  2400       2410       2420       2430       2440       2450
GAAGAAAGTGATCCTGAATATCGGTCCAGCAAATCAAAGCCATTGACCCTATTAGAA
GluGluSerAspProGluTyrArgSerSerLysSerLysProLeuThrLeuLeuGlu

     2460       2470       2480       2490       2500
TATAATTTAACTATGGACACTGCAAAGCTGTTTATGTCCTGCCTTCACGCCTGGGGT
TyrAsnLeuThrMetAspThrAlaLysLeuPheMetSerCysLeuHisAlaTrpGly

2510       2520       2530       2540       2550       2560
TTGAATGAAGTACTGGATGAAGTTTGCCTGGATCGCCTTGGAATGCTGAAACCCCAC
LeuAsnGluValLeuAspGluValCysLeuAspArgLeuGlyMetLeuLysProHis

 2570       2580       2590       2600       2610       2620
TGCACCGTATCGTTTGGCCTCTTGTCAAGAGGAGGCCATATGTCACTGATGCTGCCG
CysThrValSerPheGlyLeuLeuSerArgGlyGlyHisMetSerLeuMetLeuPro

    2630       2640       2650       2660       2670
GGTTATAATCAGCCTGCTTGTAAACTGTCACATGGGAAAACAGAAGTAGGAAGGAAG
GlyTyrAsnGlnProAlaCysLysLeuSerHisGlyLysThrGluValGlyArgLys

2680       2690       2700       2710       2720       2730
CTGCCAGCGTCTGAGGGAGTAGGAAAGGGAACTTACGGAGTGTCCCGTGCCGTCACC
LeuProAlaSerGluGlyValGlyLysGlyThrTyrGlyValSerArgAlaValThr

2740       2750       2760       2770       2780       2790
ACACAGCATCTCCTGTCTATCATTTCTTTGGCAAATACTTTAATGAGTATGACCAAT
ThrGlnHisLeuLeuSerIleIleSerLeuAlaAsnThrLeuMetSerMetThrAsn

   2800       2810       2820       2830       2840       2850
GCAACTTTTATTGGTGATCATATGAAGAAGGGTCCTACCAGGCCACCTAGACCAAGC
AlaThrPheIleGlyAspHisMetLysLysGlyProThrArgProProArgProSer

      2860       2870       2880       2890       2900
ACCCCAGACCTTTCTAAGGCAAGGGGTTCCCCTCCAACTTCCAGTAATATTGTGCAA
ThrProAspLeuSerLysAlaArgGlySerProProThrSerSerAsnIleValGln
```

```
                2910       2920       2930       2940       2950       2960
                GGACAGATTAAACAAGTTGCTGCACCTGTCGTTTCCGCTCGGTCTGATGCTGATCAC
                GlyGlnIleLysGlnValAlaAlaProValValSerAlaArgSerAspAlaAspHis

                  2970       2980       2990       3000       3010       3020
                TCTGGCTCTGACCCTCCTTCTGCTCCTGCTTTACATACCTGTTTCTTAGTAAATGAA
                SerGlySerAspProProSerAlaProAlaLeuHisThrCysPheLeuValAsnGlu

                     3030       3040       3050       3060       3070
                GGTTGGAGTCAGTTAGCTGCTATGCACTGTGTTATGCTGCCAGACCTACTGGGATTG
                GlyTrpSerGlnLeuAlaAlaMetHisCysValMetLeuProAspLeuLeuGlyLeu

               3080       3090       3100       3110       3120       3130
                GATAAATTTAGGCCTCCCCTTCTGGAGATGCTGGCCCGAAGATGGCAAGATCGATGC
                AspLysPheArgProProLeuLeuGluMetLeuAlaArgArgTrpGlnAspArgCys

                 3140       3150       3160       3170       3180       3190
                TTGGAGGTGAGAGAAGCCGCACAGGCCCTGCTTCTGGCGGAACTGAGAAGAATTGAG
                LeuGluValArgGluAlaAlaGlnAlaLeuLeuLeuAlaGluLeuArgArgIleGlu

                     3200       3210       3220       3230       3240
                CAGGCAGGCAGGAAGGAAGCCATTGATGCCTGGGCTCCTTACTTACCTCAGTACATA
                GlnAlaGlyArgLysGluAlaIleAspAlaTrpAlaProTyrLeuProGlnTyrIle

           3250       3260       3270       3280       3290       3300
                GACCACGTCATATCACCTGGAGTCACATCAGAAGCCGCGCAGACTATCACCACGGCT
                AspHisValIleSerProGlyValThrSerGluAlaAlaGlnThrIleThrThrAla

                3310       3320       3330       3340       3350       3360
                CCTGATGCCTCAGGGCCTGAAGCAAAAGTCCAGGAGGAAGAGCATGACCTTGTTGAC
                ProAspAlaSerGlyProGluAlaLysValGlnGluGluGluHisAspLeuValAsp

                   3370       3380       3390       3400       3410       3420
                GATGACATCACCACTGGTTGCTTATCAAGTGTCCCACAAATGAAAAAAATTTCTACA
                AspAspIleThrThrGlyCysLeuSerSerValProGlnMetLysLysIleSerThr

                     3430       3440       3450       3460       3470
                TCTTACGAGGAAAGACGGAAGCAAGCTACCGCTATTGTTTTACTTGGAGTAATAGGA
                SerTyrGluGluArgArgLysGlnAlaThrAlaIleValLeuLeuGlyValIleGly

               3480       3490       3500       3510       3520       3530
                GCTGAATTTGGTGCTGAAATTGAACCTCCTAAACTATTGACCAGACCTCGAAGCTCT
                AlaGluPheGlyAlaGluIleGluProProLysLeuLeuThrArgProArgSerSer

                  3540       3550       3560       3570       3580       3590
                AGCCAAATTCCTGAGGGATTCGGGTTGACTAGTGGTGGATCCAACTACTCGCTGGCC
                SerGlnIleProGluGlyPheGlyLeuThrSerGlyGlySerAsnTyrSerLeuAla

                     3600       3610       3620       3630       3640
                AGACATACTTGCAAGGCACTGACGTTTCTTCTGCTACAGCCTCCAAGCCCCAAACTT
                ArgHisThrCysLysAlaLeuThrPheLeuLeuLeuGlnProProSerProLysLeu

               3650       3660       3670       3680       3690       3700
                CCTCCACACAGCACTATCCGAAGAACAGCCATTGATCTGATTGGACGTGGGTTCACT
                ProProHisSerThrIleArgArgThrAlaIleAspLeuIleGlyArgGlyPheThr

                 3710       3720       3730       3740       3750       3760
                GTTTGGGAGCCTTACATGGATGTGTCCGCTGTTCTGATGGGGCTTCTCGAACTTTGT
                ValTrpGluProTyrMetAspValSerAlaValLeuMetGlyLeuLeuGluLeuCys

                     3770       3780       3790       3800       3810
                GCCGATGCCGAGAAACAACTTGCCAACATCACAATGGGGTTGCCTCTGAGCCCAGCA
                AlaAspAlaGluLysGlnLeuAlaAsnIleThrMetGlyLeuProLeuSerProAla
```

```
3820      3830      3840      3850      3860      3870
   GCTGACTCGGCCCGCTCTGCGAGGCATGCCCTCTCGCTCATTGCCACCGCCAGACCA
   AlaAspSerAlaArgSerAlaArgHisAlaLeuSerLeuIleAlaThrAlaArgPro

   3880      3890      3900      3910      3920      3930
   CCCGCCTTCATCACCACCATAGCCAAAGAGGTACACAGACATACGGCTCTTGCAGCA
   ProAlaPheIleThrThrIleAlaLysGluValHisArgHisThrAlaLeuAlaAla

      3940      3950      3960      3970      3980      3990
   AATACCCAATCACAGCAGAATATGCACACAACAACTCTTGCACGAGCTAAAGGGGAA
   AsnThrGlnSerGlnGlnAsnMetHisThrThrThrLeuAlaArgAlaLysGlyGlu

         4000      4010      4020      4030      4040
   ATTTTGAGAGTCATTGAAATTCTTATTGAAAAGATGCCCACAGATGTTGTGGATCTT
   IleLeuArgValIleGluIleLeuIleGluLysMetProThrAspValValAspLeu

      4050      4060      4070      4080      4090      4100
   CTCGTGGAGGTTATGGACATCATTATGTACTGCCTTGAAGGATCTTTAGTTAAAAAG
   LeuValGluValMetAspIleIleMetTyrCysLeuGluGlySerLeuValLysLys

        4110      4120      4130      4140      4150      4160
   AAAGGTCTTCAAGAATGTTTCCCAGCCATCTGCAGGTTCTACATGGTCAGCTATTAT
   LysGlyLeuGlnGluCysPheProAlaIleCysArgPheTyrMetValSerTyrTyr

         4170      4180      4190      4200      4210
   GAGCGGAATCACAGAATAGCAGTTGGAGCTCGCCATGGTTCAGTGGCCCTGTACGAC
   GluArgAsnHisArgIleAlaValGlyAlaArgHisGlySerValAlaLeuTyrAsp

   4220      4230      4240      4250      4260      4270
   ATCCGGACTGGAAAATGTCAGACAATCCATGGACACAAGGGACCAATCACTGCAGTG
   IleArgThrGlyLysCysGlnThrIleHisGlyHisLysGlyProIleThrAlaVal

    4280      4290      4300      4310      4320      4330
   GCTTTTGCTCCTGATGGAAGATATCTTGCCACCTACTCAAACACTGACAGCCACATT
   AlaPheAlaProAspGlyArgTyrLeuAlaThrTyrSerAsnThrAspSerHisIle

       4340      4350      4360      4370      4380
   TCTTTTTGGCAGATGAACACGTCACTGCTGGGAAGCATCGGCATGCTGAACTCGGCA
   SerPheTrpGlnMetAsnThrSerLeuLeuGlySerIleGlyMetLeuAsnSerAla

4390      4400      4410      4420      4430      4440
   CCTCAGCTGCGCTGCATTAAAACCTACCAGGTGCCCCCTGTGCAGCCCGCGTCCCCC
   ProGlnLeuArgCysIleLysThrTyrGlnValProProValGlnProAlaSerPro

   4450      4460      4470      4480      4490      4500
   GGCTCCCACAATGCCCTCAAGCTGGCCCGGCTCATCTGGACTTCCAACCGCAACGTC
   GlySerHisAsnAlaLeuLysLeuAlaArgLeuIleTrpThrSerAsnArgAsnVal

      4510      4520      4530      4540
   ATCCTCATGGCCCATGACGGGAAGGAGCACCGCTTCATGGTCTAA
   IleLeuMetAlaHisAspGlyLysGluHisArgPheMetValSTP
```

Table 3 shows the cDNA (SEQ ID NO: 5) and amino acid (SEQ ID NO: 6) sequences of mouse TRAG from brain as determined by double strand sequence analysis of PCR products amplified from mouse liver and brain tissue cDNA. The WD repeat elements are indicated by the underline, and the two tyrosine phosphorylation motifs are indicated in boldface type.

```
          10         20         30         40         50
TGAAGATTTGACAGGTTTGAAAACGCCATGGCAGGAAACAGCCTAGTTCTGCCCATT
                                  MetAlaGlyAsnSerLeuValLeuProIle

 60         70         80         90        100        110
GTTCTTTGGGGCCGCAAAGCACCCACACATTGCATTTCGTCAATACTGTTGACAGAT
ValLeuTrpGlyArgLysAlaProThrHisCysIleSerSerIleLeuLeuThrAsp

   120        130        140        150        160        170
GATGGGGGCACAATTGTAACTGGATGCCACGATGGACAAATATGTCTCTGGGATGTT
AspGlyGlyThrIleValThrGlyCysHisAspGlyGlnIleCysLeuTrpAspVal

       180        190        200        210        220
TCGGTAGAACTAGAAGTTAATCCCCGAGCACTGTTATTTGGCCACACAGCATCCATC
SerValGluLeuGluValAsnProArgAlaLeuLeuPheGlyHisThrAlaSerIle

230        240        250        260        270        280
ACTTGTTTGTCAAAAGCCTGCGCTTCTGGGGACAAGCGGTACACTGTGAGCGCGTCT
ThrCysLeuSerLysAlaCysAlaSerGlyAspLysArgTyrThrValSerAlaSer

  290        300        310        320        330        340
GCAAACGGAGAGATGTGCCTCTGGGATGTGAACGATGGCAGATGTATTGAATTTACC
AlaAsnGlyGluMetCysLeuTrpAspValAsnAspGlyArgCysIleGluPheThr

      350        360        370        380        390
AAGTTAGCCTGCACACACACTGGCATACAGTTCTACCAGTTCTCTGTTGGGAACCAG
LysLeuAlaCysThrHisThrGlyIleGlnPheTyrGlnPheSerValGlyAsnGln

400        410        420        430        440        450
CAAGAGGGCAGGCTCCTCTGCCATGGACATTACCCTGAAATCCTCGTTGTGGATGCC
GlnGluGlyArgLeuLeuCysHisGlyHisTyrProGluIleLeuValValAspAla

 460        470        480        490        500        510
ACCAGCCTTGAGGTGTTGTATTCCTTGGTATCGAAGATCTCTCCAGACTGGATCAGC
ThrSerLeuGluValLeuTyrSerLeuValSerLysIleSerProAspTrpIleSer

     520        530        540        550        560        570
TCCATGAGCATCATCCACTCTCAGCGGACACAAGAGGACACTGTGGTGGCGCTCTCT
SerMetSerIleIleHisSerGlnArgThrGlnGluAspThrValValAlaLeuSer

         580        590        600        610        620
GTGACAGGTATTCTGAAGGTGTGGATTGTCACCTCTGAAATGAGTGGAATGCAGGAT
ValThrGlyIleLeuLysValTrpIleValThrSerGluMetSerGlyMetGlnAsp

630        640        650        660        670        680
ACTGAGCCAATATTTGAGGAGGAATCCAAACCAATTTATTGTCAGAATTGCCAAAGC
ThrGluProIlePheGluGluGluSerLysProIleTyrCysGlnAsnCysGlnSer

   690        700        710        720        730        740
ATCTCTTTTTGTGCATTCACACAGAGGTCACTTTTGGTCGTATGCTCCAAATACTGG
IleSerPheCysAlaPheThrGlnArgSerLeuLeuValValCysSerLysTyrTrp

       750        760        770        780        790
AGGGTGTTTGATGCTGGCGACTACTCTCTGTTGTGCTCAGGTCCTAGTGAAAATGGA
ArgValPheAspAlaGlyAspTyrSerLeuLeuCysSerGlyProSerGluAsnGly

800        810        820        830        840        850
CAGACATGGACTGGAGGGGACTTTGTGTCTGCAGACAAAGTCATCATCTGGACAGAA
GlnThrTrpThrGlyGlyAspPheValSerAlaAspLysValIleIleTrpThrGlu

  860        870        880        890        900        910
AATGGGCAGAGTTACATCTACAAACTCCCTGCCAGTTGCCTTCCAGCTAGTGATTCA
AsnGlyGlnSerTyrIleTyrLysLeuProAlaSerCysLeuProAlaSerAspSer
```

```
          920        930        940        950        960
TTCCGCAGCGACGTGGGGAAAGCAGTGGAAAATTTGATCCCTCCTGTGCAGCATAGC
PheArgSerAspValGlyLysAlaValGluAsnLeuIleProProValGlnHisSer

970        980        990       1000       1010       1020
CTCTTGGATCAGAAAGATAAAGAGTTGGTAATTTGTCCTCCTGTTACTCGGTTTTTC
LeuLeuAspGlnLysAspLysGluLeuValIleCysProProValThrArgPhePhe

1030       1040       1050       1060       1070       1080
TACGGATGCAAGGAATATTTGCATAAGCTACTAATTCAGGGCGATTCTTCTGGAAGG
TyrGlyCysLysGluTyrLeuHisLysLeuLeuIleGlnGlyAspSerSerGlyArg

   1090       1100       1110       1120       1130       1140
TTAAATATCTGGAACATAGCAGACATAGCAGAGAAACAGGAAGCCGATGAAGGGCTA
LeuAsnIleTrpAsnIleAlaAspIleAlaGluLysGlnGluAlaAspGluGlyLeu

         1150       1160       1170       1180       1190
AAGATGACAACTTGTATTAGCTTGCAAGAGGCATTTGACAAGCTGAAGCCTTGCCCT
LysMetThrThrCysIleSerLeuGlnGluAlaPheAspLysLeuLysProCysPro

1200       1210       1220       1230       1240       1250
GCTGGAATTATCGATCAGCTAAGTGTGATTCCCAACAGCAATGAACCTCTTAAAGTA
AlaGlyIleIleAspGlnLeuSerValIleProAsnSerAsnGluProLeuLysVal

  1260       1270       1280       1290       1300       1310
ACTGCGAGTGTCTACATACCAGCACACGGGCGACTTGTTTGTGGCCGGGAAGATGGA
ThrAlaSerValTyrIleProAlaHisGlyArgLeuValCysGlyArgGluAspGly

       1320       1330       1340       1350       1360
AGCATCATTATTGTGCCTGCCACCCAGACGGCCATAGTTCAGCTATTGCAAGGAGAA
SerIleIleIleValProAlaThrGlnThrAlaIleValGlnLeuLeuGlnGlyGlu

1370       1380       1390       1400       1410       1420
CACATGCTCAGAAGAGGTTGGCCCCCTCACAGAACCCTCCGTGGCCATCGGAACAAA
HisMetLeuArgArgGlyTrpProProHisArgThrLeuArgGlyHisArgAsnLys

  1430       1440       1450       1460       1470       1480
GTCACGTGCCTGCTGTATCCCCATCAGGTCTCAGCTCGGTATGACCAAAGATACCTG
ValThrCysLeuLeuTyrProHisGlnValSerAlaArgTyrAspGlnArgTyrLeu

     1490       1500       1510       1520       1530
ATATCTGGAGGTGTGGATTTTTCCGTCATAATTTGGGACATATTTTCTGGAGAAATG
IleSerGlyGlyValAspPheSerValIleIleTrpAspIlePheSerGlyGluMet

1540       1550       1560       1570       1580       1590
AAACATATCTTCTGTGTTCATGGTGGTGAGATCACACAACTTCTGGTCCCCCCGGAA
LysHisIlePheCysValHisGlyGlyGluIleThrGlnLeuLeuValProProGlu

1600       1610       1620       1630       1640       1650
AACTGTAGTGCAAGAGTTCAACACTGCATCTGTTCCGTAGCCAGTGACCACTCTGTA
AsnCysSerAlaArgValGlnHisCysIleCysSerValAlaSerAspHisSerVal

   1660       1670       1680       1690       1700       1710
GGCCTCCTAAGTCTGCGAGAGAAAAAATGCATCATGTTGGCGTCTCGCCACCTCTTT
GlyLeuLeuSerLeuArgGluLysLysCysIleMetLeuAlaSerArgHisLeuPhe

         1720       1730       1740       1750       1760
CCTATTCAGGTGATCAAGTGGAGGCCTTCCGACGACTACCTGGTGGTGGGATGCACA
ProIleGlnValIleLysTrpArgProSerAspAspTyrLeuValValGlyCysThr

1770       1780       1790       1800       1810       1820
GACGGCTCTGTGTATGTCTGGCAGATGGACACTGGTGCGCTGGATCGTTGTGCAATG
```

AspGlySerValTyrValTrpGlnMetAspThrGlyAlaLeuAspArgCysAlaMet

1830 1840 1850 1860 1870 1880
GGAATCACAGCAGTGGAGATACTGAACGCTTGTGACGAAGCCGTTCCTGCCGCTGTA
GlyIleThrAlaValGluIleLeuAsnAlaCysAspGluAlaValProAlaAlaVal

1890 1900 1910 1920 1930
GACTCTCTTAGTCATCCAGCAGTCAACCTGAAACAAGCCATGACAAGACGGAGTCTC
AspSerLeuSerHisProAlaValAsnLeuLysGlnAlaMetThrArgArgSerLeu

1940 1950 1960 1970 1980 1990
GCTGCCCTTAAAAATATGGCCCACCATAAGCTGCAAACCCTTGCAACGAACCTTTTG
AlaAlaLeuLysAsnMetAlaHisHisLysLeuGlnThrLeuAlaThrAsnLeuLeu

2000 2010 2020 2030 2040 2050
GCTTCTGAGGCCTCTGACAAGGGTAATTTACCTAAATATTCCCATAACTCCCTGATG
AlaSerGluAlaSerAspLysGlyAsnLeuProLysTyrSerHisAsnSerLeuMet

2060 2070 2080 2090 2100
GTTCAAGCAATAAAGACAAACCTAACAGACCCGGATATCCATGTGCTTTTCTTTGAT
ValGlnAlaIleLysThrAsnLeuThrAspProAspIleHisValLeuPhePheAsp

2110 2120 2130 2140 2150 2160
GTGGAAGCTTTGATTATTCAACTCCTGACTGAAGAAGCCTCTAGGCCGAATACTGCA
ValGluAlaLeuIleIleGlnLeuLeuThrGluGluAlaSerArgProAsnThrAla

2170 2180 2190 2200 2210 2220
CTTATTTCCCCAGAGAATCTGCAGAAAGCATCTGGCAGTTCAGACAAAGGGGGCTCT
LeuIleSerProGluAsnLeuGlnLysAlaSerGlySerSerAspLysGlyGlySer

2230 2240 2250 2260 2270 2280
TTCCTAACTGGGAAGCGAGCAGCGGTTCTTTTCCAGCAAGTGAAAGAAACTATCAAA
PheLeuThrGlyLysArgAlaAlaValLeuPheGlnGlnValLysGluThrIleLys

2290 2300 2310 2320 2330
GAGAACATAAAGGAACACCTTCTGGATGAGGAGGAGGACGAGGAGGAGGCGAGGAGG
GluAsnIleLysGluHisLeuLeuAspGluGluGluAspGluGluGluAlaArgArg

2340 2350 2360 2370 2380 2390
CAGAGC**AGGGAGGATAGTGACCCTGAGTAC**CGGGCCAGCAAGTCCAAGCCACTGACC
GlnSer**ArgGluAspSerAspProGluTyr**ArgAlaSerLysSerLysProLeuThr

2400 2410 2420 2430 2440 2450
CTACTAGAATACAACCTTACTATGGATACCGCAAAATTATTCATGTCCTGTCTCCAC
LeuLeuGluTyrAsnLeuThrMetAspThrAlaLysLeuPheMetSerCysLeuHis

2460 2470 2480 2490 2500
GCCTGGGGCTTGAATGAAGTTCTGGATGAAGTTTGTCTCGATCGCCTCGGAATGCTG
AlaTrpGlyLeuAsnGluValLeuAspGluValCysLeuAspArgLeuGlyMetLeu

2510 2520 2530 2540 2550 2560
AAACCGCACTGCACGGTGTCCTTTGGTCTCCTATCGAGAGGAGGTCATATGTCGCTG
LysProHisCysThrValSerPheGlyLeuLeuSerArgGlyGlyHisMetSerLeu

2570 2580 2590 2600 2610 2620
ATGCTTCCCGGTTATAATCAGGCTGCTGGAAAACTGCTGCATGCGAAAGCAGAAGTA
MetLeuProGlyTyrAsnGlnAlaAlaGlyLysLeuLeuHisAlaLysAlaGluVal

2630 2640 2650 2660 2670
GGAAGGAAGCTGCCAGCAGCGGAGGGCGTAGGCAAGGGAACCTACACAGTGTCTCGA
GlyArgLysLeuProAlaAlaGluGlyValGlyLysGlyThrTyrThrValSerArg

2680 2690 2700 2710 2720 2730

```
GCGGTCACCACTCAGCATCTGTTGTCCATCATATCCCTGGCGAATACTTTAATGAGT
AlaValThrThrGlnHisLeuLeuSerIleIleSerLeuAlaAsnThrLeuMetSer

2740       2750       2760       2770       2780       2790
ATGACCAATGCAACTTTCATTGGCGATCACATGAAGAAGGGCCCCACCAGGCCGCCT
MetThrAsnAlaThrPheIleGlyAspHisMetLysLysGlyProThrArgProPro

   2800       2810       2820       2830       2840       2850
AGACCTGGCACCCCAGACCTCTCTAAGGCGAGGGATTCCCCTCCAGCCTCCAGTAAC
ArgProGlyThrProAspLeuSerLysAlaArgAspSerProProAlaSerSerAsn

      2860       2870       2880       2890       2900
ATTGTGCAAGGACAGATTAAACAAGCCGCTGCGCCTGTCGTTTCTGCTCGGTCTGAC
IleValGlnGlyGlnIleLysGlnAlaAlaAlaProValValSerAlaArgSerAsp

2910       2920       2930       2940       2950       2960
GCTGATCACTCTGGCTCTGACTCTGCCTCTCCTGCTTTACCTACCTGTTTCTTAGTA
AlaAspHisSerGlySerAspSerAlaSerProAlaLeuProThrCysPheLeuVal

  2970       2980       2990       3000       3010       3020
AATGAAGGCTGGAGCCAGTTAGCCGCCATGCACTGTGTTATGTTGCCTGACCTGCTG
AsnGluGlyTrpSerGlnLeuAlaAlaMetHisCysValMetLeuProAspLeuLeu

     3030       3040       3050       3060       3070
GGGCTGGAGAGATTCAGGCCTCCTCTCCTGGAGATGCTAGCTCGAAGATGGCAGGAC
GlyLeuGluArgPheArgProProLeuLeuGluMetLeuAlaArgArgTrpGlnAsp

3080       3090       3100       3110       3120       3130
CGATGCTTGGAGGTGAGAGAGGCTGCACAGGCCCTGCTTCTAGCTGAGCTGAGAAGG
ArgCysLeuGluValArgGluAlaAlaGlnAlaLeuLeuLeuAlaGluLeuArgArg

 3140       3150       3160       3170       3180       3190
ATTGAGCAGGCAGGACGGAAGGAGACGATCGACACCTGGGCTCCTTACTTACCTCAG
IleGluGlnAlaGlyArgLysGluThrIleAspThrTrpAlaProTyrLeuProGln

    3200       3210       3220       3230       3240
TATATGGACCATGTCATATCACCTGGCGTCACGGCGGAAGCCATGCAGACTATGGCA
TyrMetAspHisValIleSerProGlyValThrAlaGluAlaMetGlnThrMetAla

3250       3260       3270       3280       3290       3300
GCTGCTCCAGATGCCTCGGGCCCAGAAGCCAAAGTCCAGGAAGAAGAGCATGACCTG
AlaAlaProAspAlaSerGlyProGluAlaLysValGlnGluGluGluHisAspLeu

3310       3320       3330       3340       3350       3360
GTGGATGATGACATCACTGCAGGTTGCTTGTCAAGTGTCCCACAGATGAAAAAAATC
ValAspAspAspIleThrAlaGlyCysLeuSerSerValProGlnMetLysLysIle

   3370       3380       3390       3400       3410       3420
TCCACGTCTTACGAAGAAAGAAGGAAGCAGGCCACGGCTATTGTTCTCCTGGGAGTG
SerThrSerTyrGluGluArgArgLysGlnAlaThrAlaIleValLeuLeuGlyVal

      3430       3440       3450       3460       3470
ATAGGGGCAGAGTTTGGAGCTGAAATTGAACCGCCGAAGCTGCTGACCAGGCCTCGA
IleGlyAlaGluPheGlyAlaGluIleGluProProLysLeuLeuThrArgProArg

3480       3490       3500       3510       3520       3530
AGCTCCAGTCAAATCCCTGAAGGCTTTGGCTTGACCAGCGGAGGCTCCAACTACTCC
SerSerSerGlnIleProGluGlyPheGlyLeuThrSerGlyGlySerAsnTyrSer

  3540       3550       3560       3570       3580       3590
CTGGCCAGACACACGTGCAAGGCACTGACATATCTTCTGCTACAGCCACCAAGTCCC
LeuAlaArgHisThrCysLysAlaLeuThrTyrLeuLeuLeuGlnProProSerPro
```

```
          3600       3610       3620       3630       3640
AAGCTCCCTCCTCATAGCACCATCCGGAGAACTGCCACTGACCTGATTGGGCGAGGG
LysLeuProProHisSerThrIleArgArgThrAlaThrAspLeuIleGlyArgGly

3650       3660       3670       3680       3690       3700
TTCACTGTATGGGAGCCCTACATGGATGTGTCCGCTGTCCTGATGGGGCTGCTGGAG
PheThrValTrpGluProTyrMetAspValSerAlaValLeuMetGlyLeuLeuGlu

 3710       3720       3730       3740       3750       3760
CTCTGTGCAGACGCTGAGAAACAACTTGCCAACATCACAATGGGGCTGCCTCTGAGC
LeuCysAlaAspAlaGluLysGlnLeuAlaAsnIleThrMetGlyLeuProLeuSer

     3770       3780       3790       3800       3810
CCTGCAGCTGACTCTGCCCGCTCCGCAAGGCACGCCCTGTCTCTCATCGCCACAGCC
ProAlaAlaAspSerAlaArgSerAlaArgHisAlaLeuSerLeuIleAlaThrAla

3820       3830       3840       3850       3860       3870
AGACCACCCGCCTTCATCACCACCATAGCTAAGGAGGTCCACAGACACACGGCCCTT
ArgProProAlaPheIleThrThrIleAlaLysGluValHisArgHisThrAlaLeu

3880       3890       3900       3910       3920       3930
GCAGCGAACACCCAGTCCCAGCAGAGTATACACACCACCACCCTGGCGAGGGCTAAA
AlaAlaAsnThrGlnSerGlnGlnSerIleHisThrThrThrLeuAlaArgAlaLys

   3940       3950       3960       3970       3980       3990
GGCGAAATCCTGAGAGTCATTGAAATTCTCATTGAAAAGATGCCCACGGATGTGGTG
GlyGluIleLeuArgValIleGluIleLeuIleGluLysMetProThrAspValVal

        4000       4010       4020       4030       4040
GATCTTCTTGTGGAGGTGATGGATATCATCATGTACTGCCTGGAAGGATCTTTAGTG
AspLeuLeuValGluValMetAspIleIleMetTyrCysLeuGluGlySerLeuVal

4050       4060       4070       4080       4090       4100
AAAAAGAAGGGTCTTCAGGAGTGTTTTCCAGCTATCTGCAGGTTCTACATGGTCAGC
LysLysLysGlyLeuGlnGluCysPheProAlaIleCysArgPheTyrMetValSer

  4110       4120       4130       4140       4150       4160
TATTATGAGCGCAGTCACAGAATCGCAGTTGGAGCACGCCATGGCTCAGTAGCCCTG
TyrTyrGluArgSerHisArgIleAlaValGlyAlaArgHisGlySerValAlaLeu

     4170       4180       4190       4200       4210
TATGACATCCGGACTGGAAAATGTCAGACAATCCATGGGCACAAGGGACCAATCACG
TyrAspIleArgThrGlyLysCysGlnThrIleHisGlyHisLysGlyProIleThr

4220       4230       4240       4250       4260       4270
GCGGTGTCCTTTGCTCCTGATGGCCGTTACCTTGCCACCTACTCGAACACCGACAGC
AlaValSerPheAlaProAspGlyArgTyrLeuAlaThrTyrSerAsnThrAspSer

 4280       4290       4300       4310       4320       4330
CACATTTCCTTCTGGCAGATGAACACCTCACTCCTGGGAAGCATTGGAATGCTGAAC
HisIleSerPheTrpGlnMetAsnThrSerLeuLeuGlySerIleGlyMetLeuAsn

    4340       4350       4360       4370       4380
TCAGCTCCTCAGCTGCGCTGCATCAAGACCTACCAGGTCCCTCCAGTGCAGCCAGCA
SerAlaProGlnLeuArgCysIleLysThrTyrGlnValProProValGlnProAla

4390       4400       4410       4420       4430       4440
TCCCCTGGCTCCCACAACGCCCTTAAGTTGGCCCGGCTCATCTGGACTTCCAACCGG
SerProGlySerHisAsnAlaLeuLysLeuAlaArgLeuIleTrpThrSerAsnArg

4450       4460       4470       4480       4490
AATGTTATCCTCATGGCCCATGATGGGAAGGAACACCGCTTCATGGTCTGA
AsnValIleLeuMetAlaHisAspGlyLysGluHisArgPheMetValSTP
```

Table 4 shows a comparison of the amino acid sequence of the rat brain (SEQ ID NO: 2), murine brain (SEQ ID NO: 6), and human brain (SEQ ID NO: 4) TRAG proteins. The sequence specific only to brain TRAG is at positions 951 to 981 in the rat (top), 951 to 980 in the mouse (middle), and 951 to 982 in the human (bottom).

```
   1 MAGNSLVLPIVLWGRKAPTHCISSILLTDDGGTIVTGCHDGQICLWDVSVELEVNPRALL   60
   1 MAGNSLVLPIVLWGRKAPTHCISSILLTDDGGTIVTGCHDGQICLWDLSEELEVNPRALL   60
   1 MAGNSLVLPIVLWGRKAPTHCISAVLLTDDGATIVTGCHDGQICLWDLSVELQVNPRALL   60

  61 FGHTASITCLSKACASGDKRYTVSASANGEMCLWDVNDGRCIEFTKLACTHTGIQFYQFS   120
  61 FGHTAAITCLSKACASGDKQYTVSASANGEMCLWDVNDGRCIEFTKLACTHTGIQFYQFS   120
  61 FGHTASITCLSKACASSDKQYIVSASESGEMCLWDVSDGRCIEFTKLACTHTGIQFYQFS   120

 121 VGNQQEGRLLCHGHYPEILVVDATSLEVLYSLVSKISPDWISSMSIIHSQRTQEDTVVAL   180
 121 VGNQREGRLLCHGHYPEILVVDATSLEVLYSLVSKISPDWISSMSIIRSHRTQEDTVVAL   180
 121 VGNQREGRLLCHGHYPEILVVDATSLEVLYSLVSKISPDWISSMSIIRSHRTQEDTVVAL   180

 181 SVTGILKVWIVTSEMSGMQDTEPIFEEESKPIYCQNCQSISFCAFTQRSLLVVCSKYWRV   240
 181 SVTGILKVWIVTSEISGLQDTEPIFEEESKPIYCQNCQSLSFCAFTQRSLLVVCSKYWRV   240
 181 SVTGILKVWIVTSEISDMQDTEPIFEEESKPIYCQNCQSISFCAFTQRSLLVVCSKYWRV   240

 241 FDAGDYSLLCSGPSENGQTWTGGDFVSADKVIIWTENGQSYIYKLPASCLPASDSFRSDV   300
 241 FDAGDYSLLCSGPSEDGQTWTGGDFVSADKVIIWTENGQSYIYKLPASCLPASDSFRSDV   300
 241 FDAGDYSLLCSGPSENGQTWTGGDFVSSDKVIIWTENGQSYIYKLPASCLPASDSFRSDV   300

 301 GKAVENLIPPVQHSLLDQKDKELVICPPVTRFFYGCKEYLHKLLIQGDSSGRLNIWNIAD   360
 301 GKAVENLIPPVQHSLLDQKDRELVICPPVTRFFYGCKEYLHKLLIQGDSSGRLSIWNIAD   360
 301 GKAVENLIPPVQHILLDRKDKELLICPPVTRFFYGCREYFHKLLIQGDSSGRLNIWNISD   360

 361 IAEKQEADEGLKMTTCISLQEAFDKLKPCPAGIIDQLSVIPNSNEPLKVTASVYIPAHGR   420
 361 IADKQEANEGLKTTTCISLQDAFDKLKPCPAGIIDQLSVIPNSNEPLKVTASVYIPAHGR   420
 361 TADKQGSEEGLAMTTSISLQEAFDKLNPCPAGIIDQLSVIPNSNEPLKVTASVYIPAHGR   420

 421 LVCGREDGSIIIVPATQTAIVQLLQGEHMLRRGWPPHRTLRGHRNKVTCLLYPHQVSARY   480
 421 LVCGREDGSIIIVPATQTAIVQLLQGEHMLRRGWPPHRTLRGHRNKVTCLLYPHQVSARY   480
 421 LVCGREDGSIVIVPATQTAIVQLLQGEHMLRRGWPPHRTLRGHRNKVTCLLYPHQVSARY   480

 481 DQRYLISGGVDFSVIIWDIFSGEMKHIFCVHGGEITQLLVPPENCSARVQHCICSVASDH   540
 481 DQRYLISGGVDFSVIIWDIFSGEMKHIFCVHGGEITQLLVPPENCSARVQHCVCSVASDH   540
 481 DQRYLISGGVDFSVIIWDIFSGEMKHIFCVHGGEITQLLVPPENCSARVQHCICSVASDH   540

 541 SVGLLSLREKKCIMLASRHLFPIQVIKWRPSDDYLVVGCTDGSVYVWQMDTGALDRCAMG   600
 541 SVGLLSLREKKCIMLASRHLFPIQVIKWRPSDDYLVVGCTDGSVCVWQMDTGALDRCAMG   600
 541 SVGLLSLREKKCIMLASRHLFPIQVIKWRPSDDYLVVGCSDGSVYVWQMDTGALDRCVMG   600

 601 ITAVEILNACDEAVPAAVDSLSHPAVNLKQAMTRRSLAALKNMAHHKLQTLATNLLASEA   660
 601 ITAVEILNACDEAVPAAVDSLSHPAVNLKQAMTRRSLAALKNMAHHKLQTLATNLLASEA   660
 601 ITAVEILNACDEAVPAAVDSLSHPAVNLKQAMTRRSLAALKNMAHHKLQTLATNLLASEA   660

 661 SDKGNLPKYSHNSLMVQAIKTNLTDPDIHVLFFDVEALIIQLLTEEASRPNTALISPENL   720
 661 SDKGNLPKYSHNSLMVQAIKTNLTDPDIHVLFFDVEALIIQLLTEEASRPNTALISPENL   720
 661 SDKGNLPKYSHNSLMVQAIKTNLTDPDIHVLFFDVEALIIQLLTEEASRPNTALISPENL   720

 721 QKASGSSDKGGSFLTGKRAAVLFQQVKETIKENIKEHLLDEEEDEEEARRQSREDSDPEY   780
 721 QKASGSSDKGGSFLTGKRAAVLFQQVKETIKENIKEHLLDEEEDEEEVMRQRREESDPEY   780
 721 QKASGSSDKGGSFLTGKRAAVLFQQVKETIKENIKEHLLDDEEEDEEIMRQRREESDPEY   780

 781 RASKSKPLTLLEYNLTMDTAKLFMSCLHAWGLNEVLDEVCLDRLGMLKPHCTVSFGLLSR   840
```

```
 781 RASKSKPLTLLEYNLTMDTAKLFMSCLHAWGLNEVLDEVCLDRLGMLKPHCTVSFGLLSR  840
 781 RSSKSKPLTLLEYNLTMDTAKLFMSCLHAWGLNEVLDEVCLDRLGMLKPHCTVSFGLLSR  840

 841 GGHMSLMLPGYNQAAGKLLHAKAEVGRKLPAAEGVGKGTYTVSRAVTTQHLLSIISLANT  900
 841 GGHMSLMLPGYNQAAGKLLQAKAEAGRKGPATESVGKGTYTVSRAVTTQHLLSIISLANT  900
 841 GGHMSLMLPGYNQPACKLSHGKTEVGRKLPASEGVGKGTYGVSRAVTTQHLLSIISLANT  900

 901 LMSMTNATFIGDHMKKGPTRPPRPGTPDLSKARDSPPASSNIVQGQIKQAAAPVVSARSD  960
 901 LMSMTNATFIGDHMKKGPTRPPRPGTPDLSKARDSPPASSNIVQGQIKQAAAP-VSARSA  959
 901 LMSMTNATFIGDHMKKGPTRPPRPSTPDLSKARGSPPTSSNIVQGQIKQVAAPVVSARSD  960

 961 ADHSGSDS-ASPALPTCFLVNEGWSQLAAMHCVMLPDLLGLERFRPPLLEMLARRWQDRC 1019
 960 ADHSGSAS-ASPALRTCFLVNEGWSQLAAMHCVMLPDLLGLGKFRPPLLEMLARRWQDRC 1018
 961 ADHSGSDPPSAPALHTCFLVNEGWSQLAAMHCVMLPDLLGLDKFRPPLLEMLARRWQDRC 1020

1020 LEVREAAQALLLAELRRIEQAGRKETIDTWAPYLPQYMDHVISPGVTAEAMQTMAAAPDA 1079
1019 LEVREAAQALLLAELRRIEQAGRKETIDTWAPYLPQYMDHVISPGVTAEAMQTMAAAPDA 1078
1021 LEVREAAQALLLAELRRIEQAGRKEAIDAWAPYLPQYIDHVISPGVTSEAAQTITTAPDA 1080

1080 SGPEAKVQEEEHDLVDDDITAGCLSSVPQMKKISTSYEERRKQATAIVLLGVIGAEFGAE 1139
1079 SGPEAKVQEEEHDLVDDDITTGCLSSVPQMKKMSTSYEERRKQATAIVLLGVIGAEFGAE 1138
1081 SGPEAKVQEEEHDLVDDDITTGCLSSVPQMKKISTSYEERRKQATAIVLLGVIGAEFGAE 1140

1140 IEPPKLLTRPRSSSQIPEGFGLTSGGSNYSLARHTCKALTYLLLQPPSPKLPPHSTIRRT 1199
1139 IEPPKLLTRPRSSSQIPEGFGLTSGGSNYSLARHTCKALTFLLLQPPSPKLPPHSTIRRT 1198
1141 IEPPKLLTRPRSSSQIPEGFGLTSGGSNYSLARHTCKALTFLLLQPPSPKLPPHSTIRRT 1200

1200 ATDLIGRGFTVWEPYMDVSAVLMGLLELCADAEKQLANITMGLPLSPAADSARSARHALS 1259
1199 AIDLIGRGFTVWEPYMDVSAVLMGLLELCADAEKQLANITMGLPLSPAADSARSARHALS 1258
1201 AIDLIGRGFTVWEPYMDVSAVLMGLLELCADAEKQLANITMGLPLSPAADSARSARHALS 1260

1260 LIATARPPAFITTIAKEVHRHTALAANTQSQQSIHTTTLARAKGEILRVIEILIEKMPTD 1319
1259 LIATARPPAFITTIAKEVHRHTALAANTQSQQSIHTTTLARAKGEILRVIEILIEKMPTD 1318
1261 LIATARPPAFITTIAKEVHRHTALAANTQSQQNMHTTTLARAKGEILRVIEILIEKMPTD 1320

1320 VVDLLVEVMDIIMYCLEGSLVKKKGLQECFPAICRFYMVSYYERSHRIAVGARHGSVALY 1379
1319 VVDLLVEVMDIIMYCLEGSLVKKKGLQECFPAICRFYMVSYYERSHRIAVGARHGSVALY 1378
1321 VVDLLVEVMDIIMYCLEGSLVKKKGLQECFPAICRFYMVSYYERNHRIAVGARHGSVALY 1380

1380 DIRTGKCQTIHGHKGPITAVSFAPDGRYLATYSNTDSHISFWQMNTSLLGSIGMLNSAPQ 1439
1379 DIRTGKCQTIHGHKGPITAVSFAPDGRYLATYSNTDSHISFWQMNTSLLGSIGMLNSAPQ 1438
1381 DIRTGKCQTIHGHKGPITAVAFAPDGRYLATYSNTDSHISFWQMNTSLLGSIGMLNSAPQ 1440

1440 LRCIKTYQVPPVQPASPGSHNALKLARLIWTSNRNVILMAHDGKEHRFMV 1489
1439 LRCIKTYQVPPVQPASPGSHNALRLARLIWTSNRNVILMAHDGKEHRFMV 1488
1441 LRCIKTYQVPPVQPASPGSHNALKLARLIWTSNRNVILMAHDGKEHRFMV 1490
```

Table 5 shows a comparison of the amino acid sequence of murine (SEQ ID NO: 6) and Drosophila TRAG (SEQ ID NO:7) proteins. Drosophila protein sequence obtained by translating noncontiguous regions of Drosophila genomic sequence (GenBank Accession Number AL021086).

```
  1 MAGNSLVLPIVLWGRKAPTHCISSILLTDDGGTIVTGCHDGQICLWDVS-VELEVNPRAL   59
  1 MVSTNLVVPVVLWGPTAPTHCISSVFLSDDQFTLVTGCYDGQIXLWQVEPTTLKMSPRCL   60

 60 LFGHTASITCLSKACASGDKRYTVSASANGEMCLWDVNDGRCIEFTKLACTHTGIQFYQF  119
 61 LVGHSAPVLCLVRASLLPENNFLVSSSENGEMCTWDLTDGKCMEAVKLPQVHTQIQSYHT  120

120 SVGNQQEGRLLCHGHYPEILVVDATSLEVLYSLVSKISPDWISSMSIIHSQRTQEDTVVA  179
```

```
 121 --ANSEDVRLFCIGYYAEIMVMDPFSLEVIYVLSSKVKPDWISAIHVLRPMRRKDDVVLA  178

 180 LSVTGILKVWIVTSEMSGMQDTEPIFEEESKPIYCQNCQSISFCAFTQRSLLVVCSKYWR  239
 179 ITTTGTVKVWTLTGNEN--KHAEPIYENESKEIRCLNAITMNCCAQNQRTVLLVCTKYWQ  236

 240 VFDAGDYSLLCSGPSENGQTWTGGDFVSADKVIIWTENGQSYIYKLPASCLPASDSFRSD  299
 237 IYDAGDFTVLCSVIAPARERWQGGDFITSDRVMLWTDEGKGYLYKLPANCIPDNKEFHS-  295

 300 VGKAVENLIPPVQHSLLDQKDKELVICPPVTRFFYGCKEYLHKLLIQGDSSGRLNIWNIA  359
 296 --KSVVRDAPYLYYVLQHAGDK-VLSCPPAMKLLQGAGG-QHNLLR-GDSEGYISVWNVP  350

 360 DIAE---------KQEADEGLKMTTCISLQEAFDKLKPCPAGIIDQLSVIPNSNEPLKVTA  411
 351 EVPLDNISILQAKQMPPRPLKPHVCTSLVEAWSIMDPPPVGILDQLSRITES--PVKLTS  408

 412 SVYIPAHGRLVCGREDGSIIIVPATQTAIVQLLQGEHMLRRGWPPHRTLRGHRNKVTCLL  471
 409 SIYLPQQSRLVIGREDGSIVIVPATQTVMMQLLVGIKQNFSDWPSHQILYGHRGRVNCLL  468

 472 YPHQVSARYDQRYLISGGVDFSVIIWDIFSGEMKHIFCVHGGEITQLLVPPENCSARVQH  531
 469 CPSMIHSRYEKSHLLSGGIDFAVCLWDLYSGSLLHRFCVHAGEITQLLVPPESCSPRILK  528

 532 CICSVASDHSVGLLSLREKKCIMLASRHLFPIQVIKWRPSDDYLVVGCTDGSVYVWQMDT  591
 529 CICSVASDHSVTLVSLQERKCVTLASRHLFPVVTIKWAPLDDFLIVGCSDGSVYVWQMET  588

 592 GALDRCAMGITAVEILNACDEAVP--------------AAVDSLSHPAVNLKQAMTRRSL  637
 589 GHLDRVLHGMLAEEVLSACDEQAEDGGSGGGGGSNGASASEMGMANPAVHFFRGLKSRNM  648

 638 AALKNMAHHKLQTLATNLLASEASDKGN---LPKYSHNSLMVQAIKTNLTDPDIHVLFFD  694
 649 NAIR----HATQRGITQLQQLQGHNQGNFDFLMKHRSNPLVIQGLRTNPKDAESHILFFD  704

 695 VEALIIQLLTEEASRPNTALISPENLQKASGSSDKGGSFLTGKRAAVLFQQVK------E  748
 705 IEGLIFELHSEEYAQMTPATLESLGVHLQNPKDGKSMHLDASKKIGDFFNKVKNKAVDVE  764

 749 TIKENIKEHLLDEEEDEE----EARRQSREDS-----DP--EYRASKSKPLTLLEYNLTM  797
 765 KILKDKDKHGLVQKFKEKTEIVEKKVQAKVESLQKAVEPHEEQQDLKSKIASKMEVTHVM  824

 798 DTAKLFMSCLHAWGLNEVLDEVCLDRLGMLKPHCTVSFGLLSRGGHMSLMLP----GYNQ  853
 825 EVAQLLLSLLHSWGLDPHLDKMCETRLGLLRPIVPISYGVLSKAGYMSLLLPTWQNNYAI  884

 854 AAGKLLHAKAEVGRKLP---------AAEGVGKGTYTVSRAVTTQHLLSIISLANTLMSM  904
 885 PPGIQLPSSSKK-RPLPEELQRLEHLTAVFTSRLHWELSTTLTTNHILALVAMSNTLLSM  943

 905 TNATFIGDHMKKGPTRPPRPGTPDLSKARDSPPASSNIVQGQIKQAAAPVVSARSDADHS  964
 944 SAASFLPDSEK--HKKLQR-----LAQRTDS--TLSNEEE----------REELMAHHI  983

 965 GSDSASPALPTCFLVNEGWSQLAAMHCVMLPD-LLGLE--RFRPPLLEMLARRWQDRCLE 1021
 984 SQ----------IKHAWSLLATHHCFLLPDKIEALEPKKFKRPQVEMMVKRWQHHCIE 1031

1022 VREAAQALLLAELRRIEQAGRKETIDTWAPYLPQYMDHVISPGVTAEAMQTMAAAPDASG 1081
1032 IREAAQQILLGELTRMGKKGRKQLVESWAQYLPLYT-HT-EPIVGAQQQLALISQPASGG 1089

1082 PEAKVQEEEHDLVDDDIT----AGCLS----SVP-----Q----------MKKISTSYE 1117
1090 --AGSGSGGNGGVGVGVSGGGGAGSGSGPGGSVPGGDAHQDEDYEEEEEEIIRKPSSLSE 1147

1118 ERRKQATAIVLLGVIGAEFGAEI--EPP----------KLLTR-----PRSSSQIPEGF 1159
1148 LKRKQTTAVILLGVIGAEFGQDISQESPNHRGSISMATGANLTSGVAGGERRKSSVVEGF 1207

1160 GLTSGGSNYSLARHTCKALTYLLLQPPSPKLPPHSTIRRTATDLIGRGFTVWEPYMDVSA 1219
1208 GIAN-----NLARLTSMALAHLLYAPPSPKLPQYTPLRRAAIDLLGRGFTVWEPYLDVSK 1262

1220 VLMGLLELCADAEKQLANITMGLPLSPAADSARSARHALSLIATARPPAFITTIAKEVHR 1279
1263 VLLGLLEISCEG-KAVPNLNYKLPLTPQADACRTARHALRLIATARPAAFITTMAREVAR 1321

1280 HTALAANTQSQQS-IHTTTLARAKGEILRVIEILIEKMPTDVVDLLVEVMDIIMYCLEGS 1338
```

```
1322 YNTMQQNAQSINTPLTQSVLHKAKGEILQCVEMLIDKMQSEIAGLLVEVMDIALHCVDGN 1381

1339 LVKKKGLQECFPAICRFYMVSYYERSHRIAVGARHGSVALYDIRTGKCQTIHGHKGPITA 1398
1382 ELKNRGLAELCPAICKFNQISHCAQTRRIAVGANSGNLAIYELRQNKCQMIPAHTHPITS 1441

1399 VSFAPDGRYLATYSNTDSHISFWQMNTSLLGSIGMLNSAPQLRCIKTYQVPPVQPASPGS 1458
1442 LAFSPDGKYLVSYSCAENRLSFWQTSTGMFG----LGQS-QTRCTKGYSTAPIPDVS--R 1494

1459 HNALKLARLIWTSNRNVILMAHDGKEHRFMV 1489
1495 LNPMRLAKLVWINNRTVTLMLADGSETRFNV 1525
```

The Claimed Invention Is:

1. An isolated polypeptide comprising a TRAG polypeptide fragment, said fragment comprising a fragment of an amino acid sequence as shown in Table 1, 2, or 3.

2. The TRAG polypeptide fragment of claim 1, wherein said TRAG polypeptide fragment is a full-length TRAG polypeptide comprising an amino acid sequence as shown in Table 1, 2, or 3.

3. The TRAG polypeptide fragment of claim 1, further comprising a tyrosine phosphorylation motif.

4. The TRAG polypeptide fragment of claim 1, further comprising a WD-repeat-element motif.

5. The TRAG polypeptide fragment of claim 1 joined to a detectable label.

6. The TRAG polypeptide fragment of claim 5, wherein the detectable label includes a radioactive isotope, an enzyme, a chromophore, or a mixture thereof.

7. An isolated nucleic acid sequence encoding a TRAG polypeptide fragment, said TRAG fragment comprising a fragment of an amino acid sequence as shown in Table 1, 2, or 3.

8. The nucleic acid sequence of claim 7, comprising a nucleotide sequence coding for an amino acid sequence as shown in Table 1, 2, or 3.

9. The nucleic acid sequence of claim 7, wherein the nucleic acid sequence is codon optimized for a specific host cell.

10. The nucleic acid sequence of claim 7 joined to a detectable label.

11. The nucleic acid sequence of claim 7, wherein said nucleic acid sequence is DNA.

12. The nucleic acid sequence of claim 11, wherein the DNA is cDNA.

13. The nucleic acid sequence of claim 7, wherein the nucleic acid sequence is RNA.

14. The nucleic acid sequence of claim 13, wherein the RNA is mRNA.

15. A vector comprising a polynucleotide encoding the TRAG polypeptide fragment of claim 1.

16. The vector of claim 15, wherein the nucleic acid is operably linked to at least one control sequence capable of being recognized by a host cell transformed with the vector.

17. A host cell comprising the vector of claim 16.

18. A process for producing TRAG polypeptide fragments comprising culturing the host cell of claim 17 under conditions such that the TRAG polypeptide fragment is produced.

19. A TRAG polypeptide fragment produced by the method of claim 18.

20. A TRAG antisense oligonucleotide comprising a nucleotide sequence that is complementary to an mRNA encoding a polypeptide comprising a TRAG polypeptide fragment of claim 1.

21. A chimeric molecule comprising a TRAG polypeptide fragment of claim 1 fused to a heterologous amino acid sequence.

22. An isolated TRAG specific polypeptide comprising an F_{ab} fragment from an antibody capable of specifically binding to a TRAG polypeptide fragment of claim 1.

23. The isolated TRAG specific polypeptide of claim 22, wherein said polypeptide comprises an isolated antibody.

24. The TRAG specific polypeptide of claim 23, wherein said antibody is a polyclonal, monoclonal, or chimeric antibody.

25. A method of assaying a sample for a polynucleotide encoding a TRAG polypeptide fragment comprising detecting the presence or absence of said polynucleotide in said sample utilizing a nucleic acid probe capable of hybridizing with the nucleic acid sequence of claim 7.

26. A method of assaying a sample for a TRAG polypeptide fragment of claim 1 comprising detecting the presence or absence of said TRAG polypeptide fragment in said sample utilizing an isolated TRAG specific polypeptide, said TRAG specific polypeptide comprising a F_{ab} fragment from an antibody capable of specifically binding to the TRAG polypeptide fragment.

27. A method of reducing expression of a TRAG polypeptide fragment of claim 1 in a cell comprising exposing the cell to an oligonucleotide of at least about 15 nucleotides which are complementary to a TRAG mRNA.

28. A method for producing cell lines having an altered phenotype comprising:

(i) transfecting in vitro mammalian cells with a DNA vector encoding a TRAG polypeptide fragment of claim 1;

(ii) expressing the TRAG polypeptide fragment in said cells; and

(iii) selecting for cells having an altered phenotype.

FIG. 1A

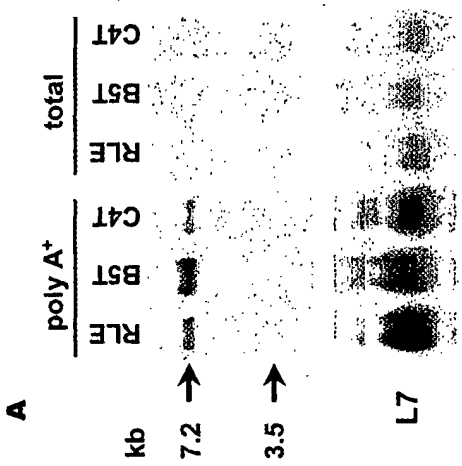
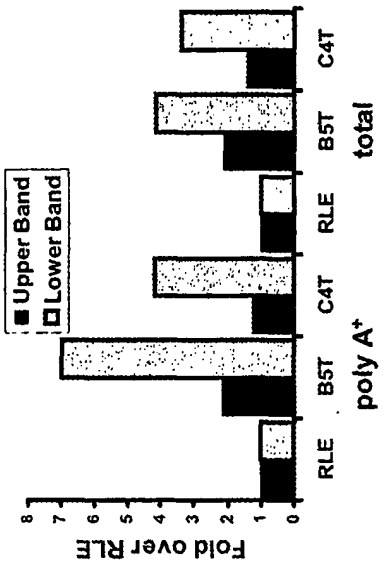
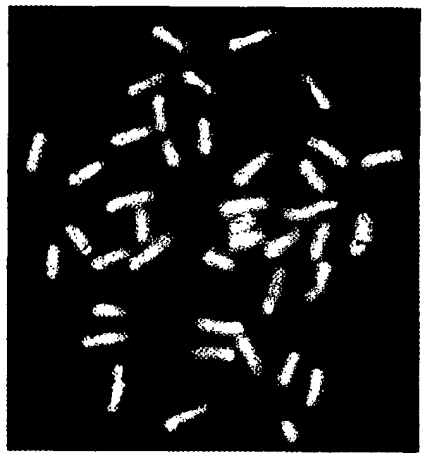
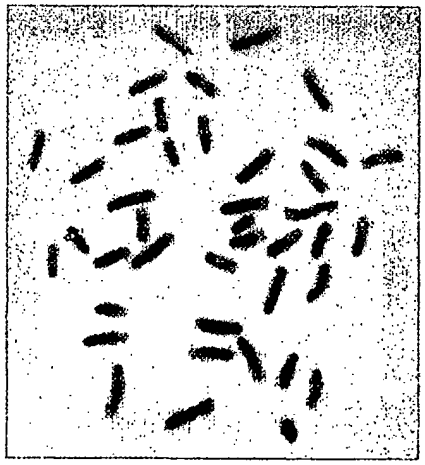

A
poly A+
total
RLE
B5T
C4T
RLE
B5T
C4T
kb
7.2
3.5
L7

FIG. 1B


B
Upper Band
Lower Band
Fold over RLE
0
1
2
3
4
5
6
7
8
RLE
B5T
C4T
RLE
B5T
C4T
poly A+
total



Mouse Ch. 18 paint



TRAG probe

FIG. 2A

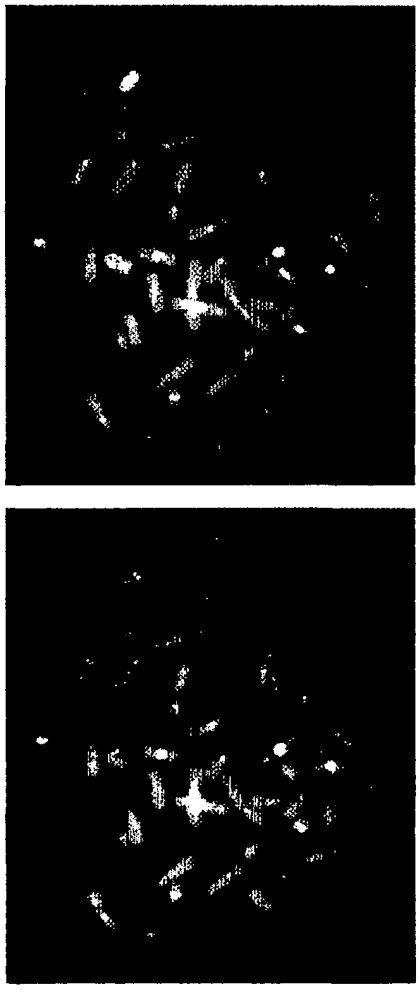


FIG. 2B

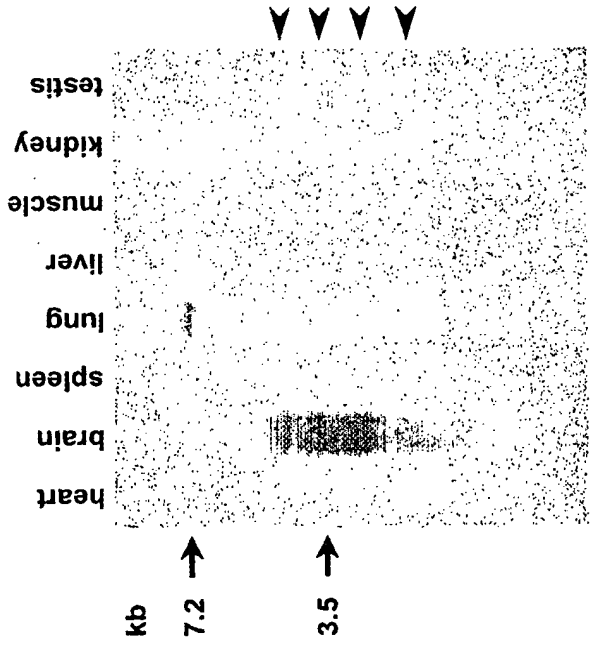


FIG. 3A

FIG. 3C

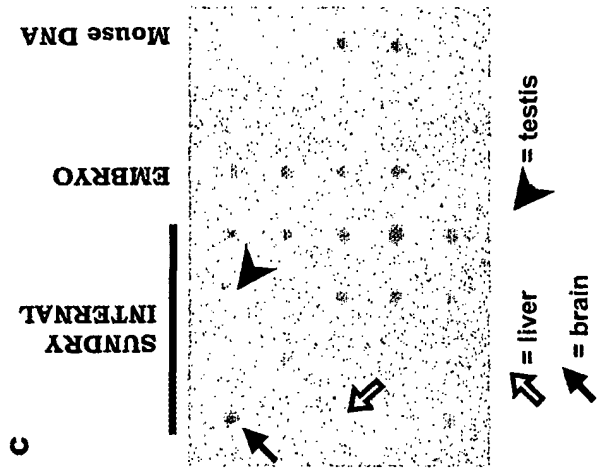

C
SUNDRY INTERNAL
EMBRYO
Mouse DNA
= testis
= liver
= brain

FIG. 3B

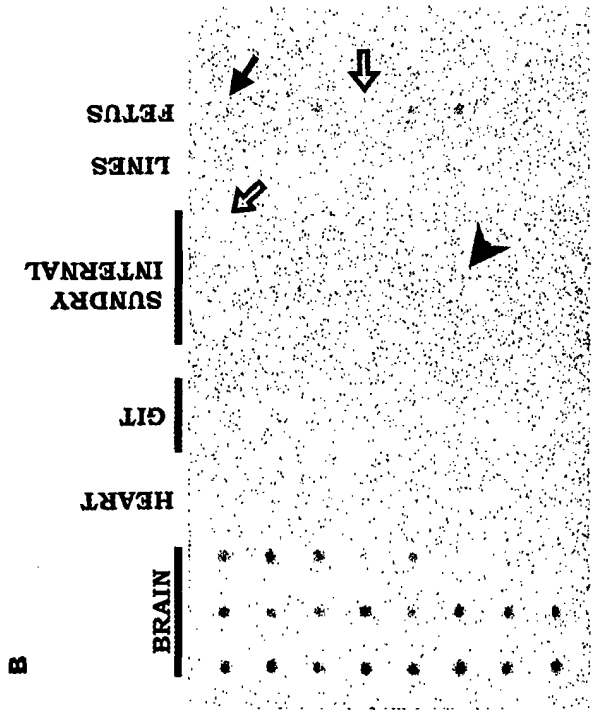

B
BRAIN
HEART
GIT
SUNDRY INTERNAL
LINES
FETUS

FIG. 4

FIG. 5B

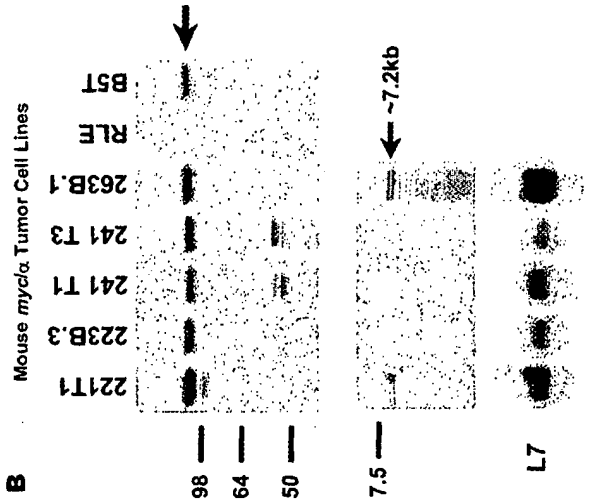
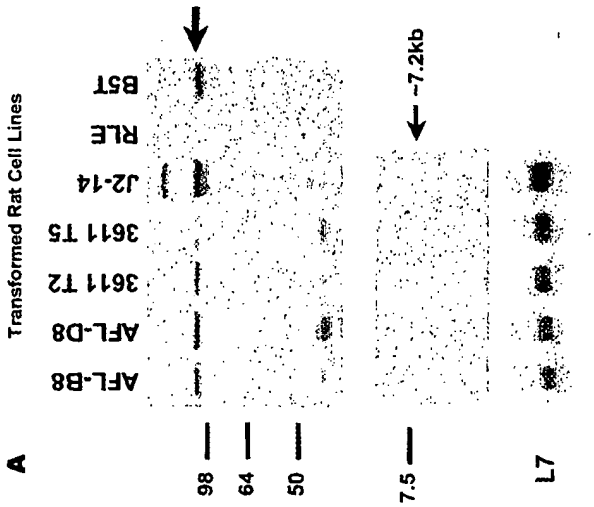

B
Mouse myc/α Tumor Cell Lines
B5T
RLE
263B.1
241 T3
241 T1
223B.3
221T1
~7.2kb
98
64
50
7.5
L7

FIG. 5A


A
Transformed Rat Cell Lines
B5T
RLE
J2-14
3611 T5
3611 T2
AFL-D8
AFL-B8
~7.2kb
98
64
50
7.5
L7

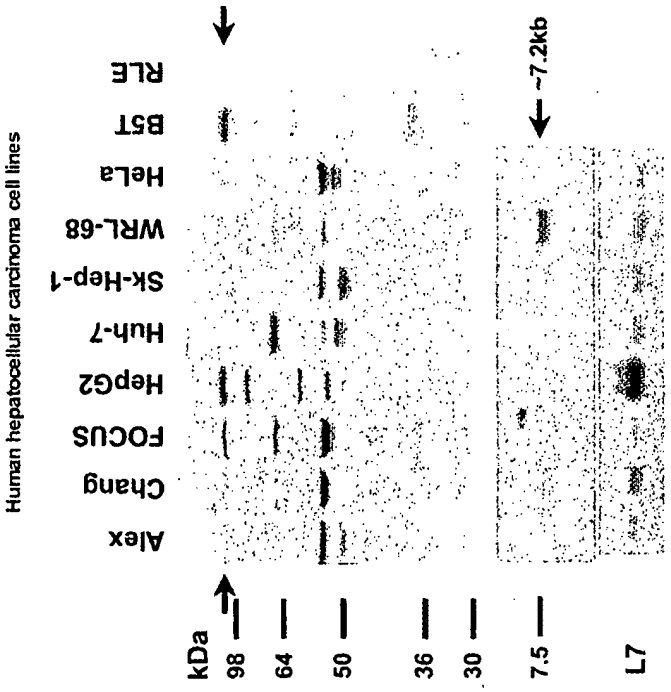

Human hepatocellular carcinoma cell lines
Alex
Chang
FOCUS
HepG2
Huh-7
Sk-Hep-1
WRL-68
HeLa
B5T
RLE
~7.2kb
kDa
98
64
50
36
30
7.5
L7

FIG. 5C

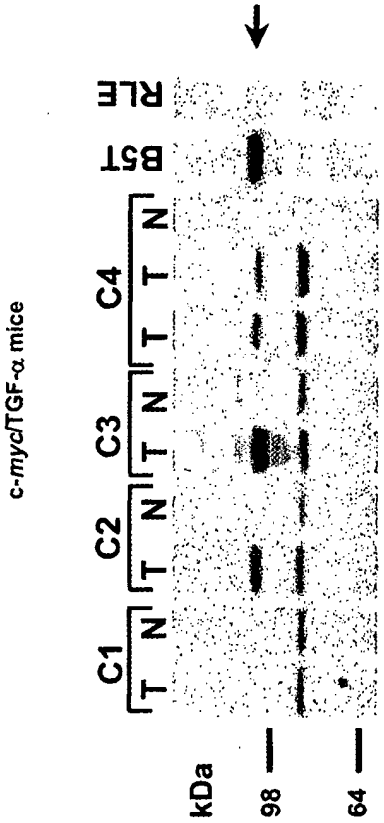


FIG. 5D

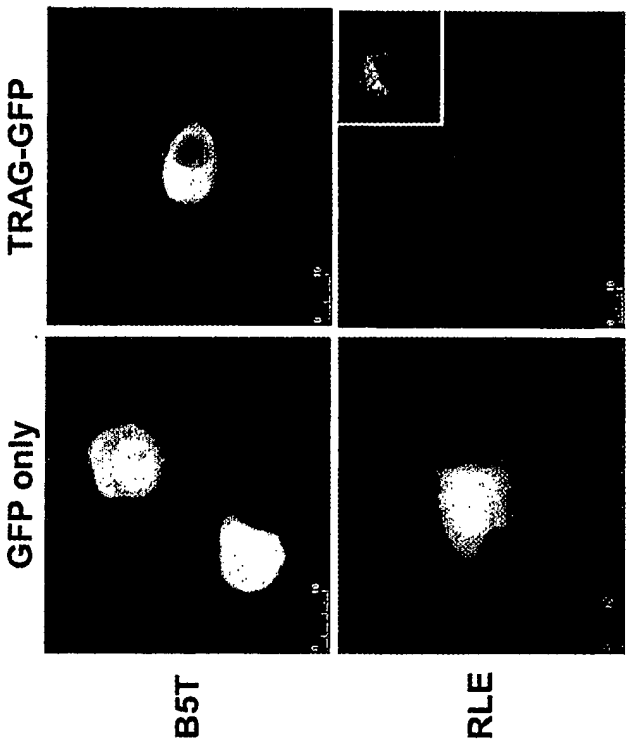


FIG. 6A

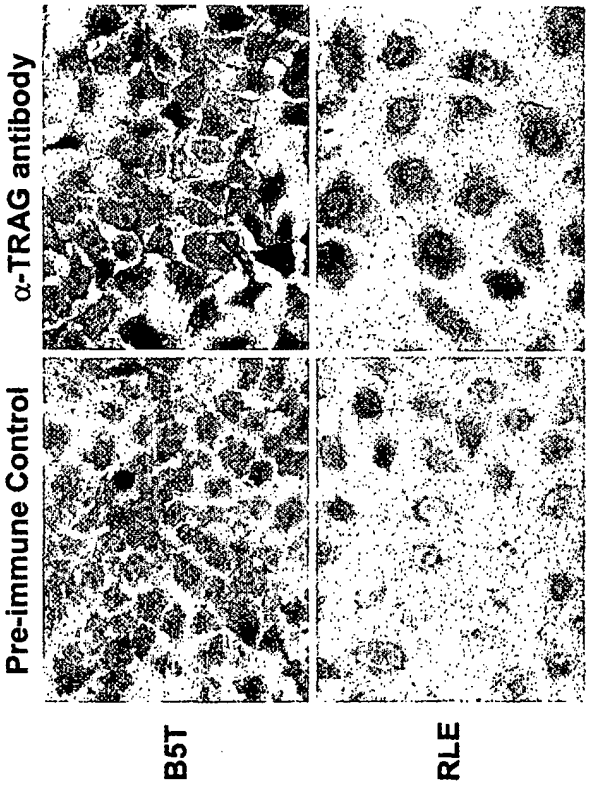


FIG. 6B

INTERNATIONAL SEARCH REPORT

Ir[illegible]ational Application No
PCT/US 01/04475

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C12N5/10 C12N15/62 C07K16/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, STRAND, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SANDERS SEAN ET AL: "TRAG: A novel gene associated with TGF-beta resistance." PROCEEDINGS OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH ANNUAL, no. 41, March 2000 (2000-03), page 357 XP001011479 91st Annual Meeting of the American Association for Cancer Research.;San Francisco, California, USA; April 01-05, 2000, March, 2000 ISSN: 0197-016X the whole document --- -/--	1-4,7,8, 10-14

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 July 2001	23/07/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Holtorf, S

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/04475

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SANDERS S ET AL: "Assignment of WDR7 (alias TRAG, TGF-beta resistance associated gene) to orthologous regions of human chromosome 18q21.1fwdarwq22 and mouse chromosome 18D.1-E.3 by fluorescence in situ hybridization." CYTOGENETICS AND CELL GENETICS, vol. 88, no. 3-4, 2000, pages 324-325, XP001011475 Fourth International Chromosome 6 Workshop;England, Cambridge, UK; June 10-12, 1999 ISSN: 0301-0171 the whole document ---	1-4,7,8, 10-14
X	NAGASE T ET AL: "PREDICTION OF THE CODING SEQUENCES OF UNIDENTIFIED HUMAN GENES. IX. THE COMPLETE SEQUENCES OF 100 NEW CDNA CLONES FROM BRAIN WHICH CAN CODE FOR LARGE PROTEINS IN VITRO" DNA RESEARCH,JP,UNIVERSAL ACADEMY PRESS, vol. 5, no. 5, 1998, pages 31-39, XP000878819 ISSN: 1340-2838 cited in the application the whole document	1-4,7,8, 11-14
X	& DATABASE EMBL SEQUENCE LIBRARY 'Online! 10 April 1998 (1998-04-10) OHARA, O., ET AL.: "Homo sapiens mRNA for KIAA0541 protein" cited in the application accession no. AB011113 ---	1-4,7,8, 11-14
X	DATABASE EMBL SEQUENCE LIBRARY 'Online! 20 September 1999 (1999-09-20) BONALDO, M.F., ET AL. : "normalization and substraction: two approaches to facilitate gene discovery" XP002171571 accession no. AW045623 ---	7
A	DATTA PRAN K ET AL: "Identification of STRAP, a novel WD domain protein in transforming growth factor-beta signaling." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 52, 25 December 1998 (1998-12-25), pages 34671-34674, XP002171570 ISSN: 0021-9258 cited in the application --- -/--	

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/04475

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CHEN R -H ET AL: "A WD-DOMAIN PROTEIN THAT IS ASSOCIATED WITH AND PHOSPHORYLATED BY THE TYPE II TGF-BETA RECEPTOR" NATURE,GB,MACMILLAN JOURNALS LTD. LONDON, vol. 377, 12 October 1995 (1995-10-12), pages 548-552, XP002071531 ISSN: 0028-0836 cited in the application	
A	WO 99 43811 A (THORGEIRSSON SNORRI S ;US HEALTH (US); ZHANG MINGHUANG (US); WOITA) 2 September 1999 (1999-09-02) the whole document	
A	NEER E ET AL: "The ancient regulatory-protein family of WD-repeat proteins" NATURE,GB,MACMILLAN JOURNALS LTD. LONDON, vol. 371, 22 September 1994 (1994-09-22), pages 297-300, XP002081318 ISSN: 0028-0836 cited in the application	
P,X	DATABASE EMBL SEQUENCE LIBRARY 'Online! 11 November 2000 (2000-11-11) SANDERS, S. AND THORGEIRSSON, S.S.: "TAG: a novel gene associated with TGF-beta resistance" XP002171572 accession no. AF305813	1-4,7,8, 11-14

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/04475

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9943811 A	02-09-1999	AU 2790599 A	15-09-1999